ENERGY-EFFICIENT DESIGN IN WIRELESS COMMUNICATIONS NETWORKS

A Dissertation
Presented to
The Academic Faculty

By

Cong Xiong

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
August 2014

ENERGY-EFFICIENT DESIGN IN WIRELESS COMMUNICATIONS NETWORKS

Approved by:

Dr. Geoffrey Ye Li, Advisor
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Chin-Hui Lee
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Gordon L. Stüber
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Xingxing Yu
Professor, School of Mathematics
Georgia Institute of Technology

Dr. John R. Barry
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: May 01, 2014

To Deji Xiong and Ping Wang - my parents.

ACKNOWLEDGEMENTS

I would like to present my gratitude to my advisor, Prof. Geoffrey Ye Li, who offered me the precious opportunity to pursue my Ph.D. study at Georgia Tech and patiently guided me through the successful fulfillment. I appreciate his care and concern for both my intellectual and personal growth. What I have learnt from him will definitely benefit my future life and career.

I am honored to have Profs. Gordon L. Stüber, John R. Barry, Chin-Hui Lee, and Xingxing Yu to be my dissertation committee members. Their broad perspective and suggestions have helped me a lot in refining this dissertation. I'm also grateful to Prof. Leonard J. Cimini, Jr. at University of Delaware. The collaboration with him and his group was quite inspiring and turned out a great joy.

I am deeply indebted to my aunt, Rong Wang, who cared me a lot since my college era. She also helped me a lot in academic writing when I was a graduate student at Beijing University of Posts and Telecommunications. Although she is always busy as a world-class chemist, she gives me a top priority.

I would also like to thank Xiangwei Zhou and Jun Ma, who have helped me a lot like brothers since I came to USA. My thanks also go to other wonderful labmates, such as Zhikun Xu, Liying Li, Jiancun Fan, Deli Jia, Hao He, and Daquan Feng, as well my dear friend Yun Wei and other friends. My special thanks belong to Lu Lu, who brought me a lot of joy and good memory during the Ph.D. study. She is such a cute girl.

Last but not the least, I would like to express my sincerest gratitude to my parents, Ping Wang and Deji Xiong. They are always proud of me. Their love, encouragement, and support are always the source of my self-motivation. This dissertation is dedicated to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The widespread application of wireless services and the requirements of ubiquitous access have recently triggered rapidly booming energy consumption in wireless communications networks. Such escalation of energy consumption in wireless networks causes high operational expenditure from electricity bills for operators, unsatisfactory user experience due to limited battery capacity of wireless devices, and a large amount of greenhouse gas emission. *Green radio* (GR), which emphasizes both *energy efficiency* (EE) and spectral efficiency (SE), has been proposed as an effective solution and is becoming the mainstream for future wireless network design. Unfortunately, EE and SE do not always coincide and may even sometimes conflict. In this dissertation, we focus on energy-efficient transmission and resource allocation techniques for *orthogonal frequency division multiple access* (OFDMA) networks and the joint energy-efficient design of OFDMA and other promising wireless communications techniques, such as *cognitive radio* (CR) and two-way relay.

Firstly, we investigate the principles of energy-efficient design for pure OFDMA networks. As the first step, we study the fundamental interrelationship between EE and SE in downlink OFDMA networks and analyze the impacts of channel gain and circuit power on the EE-SE relationship. We establish a general EE-SE optimization framework, where the overall EE, SE and per-user *quality-of-service* (QoS) are all considered. Under this framework, we find that EE is quasiconcave in SE and decreases with SE when SE is large enough. These findings are very helpful guidelines for designing energy- and spectral-efficient OFDMA. To facilitate the application of energy-efficient resource allocation, we then investigate the energy-efficient resource allocation in both downlink and uplink OFDMA networks. For the downlink transmission, the generalized EE is maximized while for the uplink case the minimum individual EE is maximized, both under prescribed per-user minimum data rate requirements. For both transmission scenarios, we first provide the

optimal solution and then develop an computationally efficient suboptimal approach by exploring the inherent structure and property of the energy-efficient design. Then we study energy-efficient design in downlink OFDMA networks with effective capacity-based delay provisioning for delay-sensitive traffic. By integrating information theory with the concept of effective capacity, we formulate and solve an EE optimization problem with statistical delay provisioning. We also analyze the tradeoff between EE and delay, the relationship between spectral-efficient and energy-efficient designs, and the impact of system parameters, including circuit power and delay exponents, on the overall performance.

Secondly, we consider joint energy-efficient design of OFDMA and CR and two-way relay, respectively, to further enhance the EE and SE of wireless networks. We study energy-efficient opportunistic spectrum access strategies for an OFDMA-based CR network with multiple *secondary users* (SUs). Both worst EE and average EE of the SUs are considered and optimized subject to constraints including maximum transmit power and maximum interference to *primary user* (PU) system. For both cases, we first find the optimal solution and then propose a low-complexity suboptimal alternative. The results show that the energy-efficient CR strategies significantly boost EE compared with the conventional spectral-efficient CR ones while the low-complexity suboptimal approaches can well balance the performance and complexity. Then we study energy-efficient resource allocation for OFDMA-based two-way relay, which aims at maximizing the aggregated EE utility while provisioning proportional fairness in EE among different terminal pairs. Different from most exist energy-efficient design, we consider a new circuit power model, where the dynamic circuit power is proportional to the number of active subcarrier. For low-complexity solution, we propose an EE-oriented sequential subchannel assignment policy and discover the sufficient condition for early termination of the sequential subchannel assignment without losing the EE optimality. It is found that the energy-efficient transmission does not necessarily make all the subcarriers active, which is another useful principle for practical energy-efficient system design.

CHAPTER 1

INTRODUCTION

1.1 Motivation

The global wireless traffic has been approximately doubling each year during the last few years [1]. The widespread application of wireless services and the requirements of ubiquitous access have triggered rapidly booming energy consumption in wireless communication networks. The escalation of energy consumption in radio access networks, which take up more than 70% of the total energy consumption in wireless communications networks [2], results in high operational expenditure for operators. The large energy consumption at mobile terminals causes unsatisfactory user experience, considering the limited battery capacity and and the slow advancement of battery technology [3]. In addition, a large amount of greenhouse gas emission is generated, as reported that wireless communications account for more than 9% of the global carbon dioxide emissions originated from *information and communications technologies* (ICT) [4]. The rising energy concerns prompt the fast development of *green radio* (GR) [5], which emphasizes *energy efficiency* (EE) and *spectral efficiency* (SE).

Orthogonal frequency division multiple access (OFDMA) [6, 7, 8] has been extensively studied and proposed for next-generation wireless communications systems, such as 3GPP LTE [9] and WiMAX [10], partially because of its high SE. *Cognitive radio* (CR) [11, 12, 13, 14], which enables opportunistic utilization of largely underutilized licensed spectrum, has also been intensively investigated over the past two decades. Meanwhile, two-way relay [15, 16, 17, 18] proves to be more spectral-efficient than direct transmission and one-way relay in many typical scenarios. All these techniques help to increase the SE and throughput of wireless communications. However, EE performance and energy-efficient design of them are seldom investigated in literature. To keep pace with GR, it is necessary for them to guarantee a certain level of EE and well balance EE and SE.

On the one hand, to evolve toward the new energy-efficient regime of wireless communications, it is of great significance to figure out the principles of energy-efficient transmission in pure OFDMA networks, including the following aspects.

- For point-to-point transmission in an *additive white Gaussian noise* (AWGN) channel, the relation between EE and SE is either a cup shape curve without considering circuit power or a bell shape curve if circuit power is considered [5]. However, EE-SE interrelationship for OFDMA networks is much more complicated and not clear. Moreover, the impacts of system parameters such as channel gain and circuit power on the interrelationship are worth investigating.

- To implement energy-efficient OFDMA, it is necessary to develop optimal and low-complexity suboptimal resource allocation algorithms for both the downlink and uplink of OFDMA networks, where the per-user minimum data rate requirement and fairness/priority issues are considered.

- Delay provisioning for delay-sensitive high-data-rate services is an important and challenging issue for future wireless communication networks [19]. Therefore, how to achieve high EE while ensuring the per-user delay requirement for OFDMA networks with delay-sensitive traffic is another important topic.

On the other hand, OFDMA is usually adopted with other advance transmission techniques (e.g., CR and two-way relay) to further boost the system performance. Hence, the EE of OFDMA networks equipped these techniques should be further exploited.

- To reconcile CR and energy-efficient transmission for high spectrum utilization and EE simultaneously, it is undoubtedly of great necessity and value to study spectrum access targeting at EE of SUs in CR. Specifically, it is interesting to study how to efficiently optimize average EE and worst EE of OFDMA-based CR networks.

- Two-way relay can potentially further increase the EE and SE of OFDMA. To develop energy-efficient OFDMA-based two-way relay networks, it is desirable to concentrate on energy-efficient joint power and subchannel allocation, including active subchannel selection, while providing proportional fairness.

1.2 Literature Review

In this section, we review state-of-the-art techniques for energy- and spectral-efficient design of OFDMA networks and OFDMA-based CR and two relay networks, respectively.

1.2.1 Energy- and Spectral-Efficient Design of OFDMA Networks

1.2.1.1 Energy- and Spectral-Efficiency Tradeoff

During the past decades, EE, which is commonly defined as information bits per unit transmit energy, has been studied from the information-theoretic perspective for various scenarios [20]. For an additive white Gaussian noise (AWGN) channel, it is well known that for a given transmit power, P, and system bandwidth, B, the channel capacity is $R = \frac{1}{2}\log_2\left(1+\frac{P}{N_0 B}\right)$ bits per real dimension or degrees of freedom (DOF), where N_0 is the noise power spectral density. According to the Nyquist sampling theory, DOF per second is $2B$. Therefore, the channel capacity is $C = 2BR$ bits per second. Consequently, EE is [5, 21]

$$\eta_{EE} \triangleq \frac{C}{P} = \frac{2R}{N_0\left(2^{2R}-1\right)}. \tag{1.1}$$

From (1.1), it is obvious that η_{EE} decreases monotonically with R, with $(\eta_{EE})_{\max} = 1/\left(N_0 \ln 2\right)$ as $R \to 0$, and $(\eta_{EE})_{\min} = 0$ as $R \to \infty$. And EE and SE has a simple tradeoff relationship as shown in Figure 1.1 [5].

The EE bounds derived from the information-theoretic analysis might not be achieved in practical systems as a result of performance loss of capacity-approaching channel codes, imperfect knowledge of *channel state information* (CSI) [22], cost of synchronization [23],

and transmission associated electronic circuit energy consumption [24, 25, 26, 27, 28]. Among these factors, electronic circuit energy consumption changes the fundamental tradeoff between EE and data rate. Taking circuit energy consumption into consideration, EE is redefined as information bits per unit energy (not only transmit energy), where an additional circuit power factor, P_c, needs to be added in the denominator of (1.1). Accordingly, the η_{EE}-versus-R curve will turn from a cup shape to a bell shape as shown in Figure 1.1 from [5].

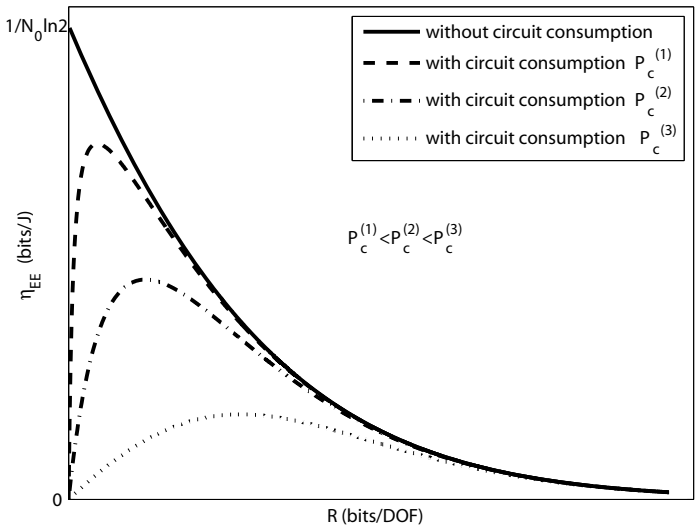


Figure 1.1. Tradeoff between EE and rate in an AWGN channel.

In [29], EE-SE tradeoff with consideration of circuit power has been studied for energy-constrained wireless multihop networks with a single source-destination pair. It reveals that the common power (on each hop) strategy yields better EE-SE tradeoff than the common rate (on each hop) one and at higher rates fewer hops are good for EE while at lower rates more hops are preferred. For two-hop AWGN relay channels, EE-SE tradeoff for different relaying strategies with different forwarding methods have been investigated and compared in [30]. It is demonstrated that at lower rate region noncooperative relaying outperforms cooperative one in EE while at higher rate region the latter is better. In [31], the tradeoff between EE and SE in downlink multiuser *distributed antenna systems* (DAS) is

revealed. In [32], a multi-criteria optimization method is further proposed to systematically investigate the relationship between EE and SE in DAS with proportional fairness among the users. In [33], for the single-user *multi-input multi-output* (MIMO) Rayleigh fading channel is characterized by a generic closed-form approximation, which exhibits a greater accuracy for a wider range of SE values and antenna configurations. Another unified metric for EE and SE tradeoff design in point to point wireless networks is proposed in [34]. Both EE and SE are simultaneously optimized through a multi-object optimization problem, where the Pareto optimal set is further characterized. In [35], the impact of *power amplifier* (PA) on the SE-EE tradeoff of *orthogonal frequency division multiplex* (OFDM) systems is analyzed. It shows that a practical PA has an SE-EE tradeoff that has a turning point and decreases sharply after its maximum EE point. However, EE-SE relation for general multi-user networks, including downlink OFDMA network, is much more complicated and not clear yet. And how to allocate system resource to tradeoff EE and SE efficiently for these networks is a nontrivial question.

1.2.1.2 Energy- and Spectral-Efficient Resource Allocation

In OFDMA networks, system resources, such as subcarriers and transmit power, need to be appropriately allocated to different UEs to achieve high performance. Two most commonly used classes of dynamic resource allocation schemes are *rate adaptation* (RA) that maximizes the throughput and *margin adaptation* (MA) that minimizes the total transmit power [36]. However, neither of them is necessarily energy-efficient evaluated under the bits-per-Joule metric [3]. Recently, more attention has been paid to energy-efficient design in OFDMA networks. As a special case of OFDMA systems, energy-efficient OFDM transmission, which maximizes the EE (i.e., bits-per-Joule) or equivalently minimizes the energy cost per bit (i.e., Joule-per-bit), has been addressed with consideration of circuit energy consumption in [26, 37, 28], respectively. It is demonstrated that a unique global maximum EE exists and can be obtained by the proposed algorithms therein. For multiuser OFDM or namely OFDMA networks, energy-efficient design is generally much more

complicated than the single-UE case. For one special uplink OFDMA transmission with flat fading channels for UEs, the energy-efficient design is considered and demonstrated to consume less energy than the traditional fixed power schemes for a fixed amount of bits [27]. In [38], long-time EE is optimized for an uplink OFDMA network with a long transmission duration. Low-complexity schedulers are derived in closed-form for both arithmetic and geometric average EE metrics by some approximation. *Time domain synchronous OFDM* (TDS-OFDM) has a higher spectrum and energy efficiency than standard cyclic prefix OFDM by replacing the unknown CP with a known *pseudorandom noise* (PN) sequence. However, due to mutual interference between the PN sequence and the OFDM data block, TDS-OFDM cannot support high-order modulation schemes such as 256-QAM in realistic static channels with large delay spread in fast fading channels. In [39], multiple *inter-block-interference* (IBI)-free regions of small size is exploited to realize simultaneous multi-channel reconstruction. The proposed OFDM has a higher spectrum and energy efficiency than CP-OFDM by more than 10% and 20% respectively in typical applications. In [40], an energy-efficient resource-allocation problem with proportional fairness for downlink multiuser OFDM systems with distributed antennas is studied. Because of the nonconvexity of the problem, a low-complexity suboptimal algorithm, which separates subcarrier allocation and power allocation, is proposed for suboptimal solution. In [41], resource allocation for energy-efficient communication in multi-cell OFDMA downlink networks with cooperative *base stations* (BSs) is studied. The resource allocation problem for joint energy-efficient BS zero-forcing beamforming (ZFBF) transmission is considered, which takes into account the circuit power consumption, the limited backhaul capacity, and the minimum required data rate. The problem in fractional form is transformed into an equivalent optimization problem in subtractive form, which enables the derivation of an efficient iterative resource allocation algorithm. In [42], energy-efficient communication in an OFDMA downlink network with a large number of transmit antennas is studied. The considered problem is modeled as a non-convex optimization problem which takes into

account the circuit power consumption, imperfect *channel state information at the transmitter* (CSIT), and different QoS requirements including a minimum required data rate and a maximum tolerable channel outage probability. The power allocation, data rate adaptation, antenna allocation, and subcarrier allocation policies are optimized for maximization of the EE of data transmission. Delay provisioning for delay-sensitive high-data-rate services is an important and challenging issue for future wireless communication networks. Due to the time-varying nature of wireless channels, it is difficult and impractical to impose a deterministic delay guarantee for services over wireless networks. To address this issue, the effective capacity [43, 44, 45, 46] has been widely adopted to provide the statistical delay provisioning to ensure a small steady-state delay violation probability. In [47], EE in fading channels in the presence of delay constrained QoS constraints is studied. SE and bit energy tradeoff is analyzed in the low-power and wideband regimes by employing the effective capacity formulation, rather than the Shannon capacity. Through this analysis, energy requirements under QoS constraints are identified. The analysis is conducted under two assumptions: perfect CSI available only at the receiver and perfect CSI available at both the receiver and transmitter. In particular, it is shown in the low-power regime that the minimum bit energy required under QoS constraints is the same as that attained when there are no such limitations. Similarly, in [45], EE of fixed-rate transmissions is further studied in the presence of queueing constraints and channel uncertainty. However, there is limited work on the energy-efficient design with delay provisioning for delay-sensitive traffic in the downlink OFDMA networks.

1.2.2 Energy- and Spectral-Efficient Design of OFDMA-based CR and OFDMA-based Two-Way Relay Networks

1.2.2.1 Energy- and Spectral-Efficient Design of OFDMA-based CR

Various secondary spectrum access criteria and schemes in CR networks have been developed for different target objectives, including interference, power, throughput, delay, and price [48]. To reconcile CR and energy-efficient transmission for high spectrum utilization

and EE simultaneously, it is undoubtedly of great necessity and value to study spectrum access targeting at EE in CR. In [49], energy-efficient sequential sensing of multiple licensed frequency channels in a single-SU CR network has been studied. In [50], the overall EE of an OFDM-based CR system is studied under the total power constraint, the interference power constraint and the rate constraint. A novel method named waterfilling factors aided search is proposed to solve the EE optimization problems with multiple constraints. In [51], the system EE of an OFDM-based CR under the consideration of many practical limitations, such as transmission power budget of the CR system, interference threshold of PUs and traffic demands of SUs, is optimized. By exploiting the structure of the problem, an efficient barrier method is developed for near-optimal solution. In [52], energy-efficient scheduling in CR, in which a *cognitive base station* (CBS) makes frequency allocations to the CRs at the beginning of each frame according to the diversity among CRs' queues and channel capacities in terms of number of bits and the channel switching cost from one frequency to another, is studied. A polynomial-time heuristic algorithm, which allocates each idle frequency to the CR that attains the highest energy efficiency at this frequency, is proposed as a low-complexity energy-efficient scheduler. However, there still is limited work on spectrum access for multiple SUs targeting at the average and/or worst EE in OFDM-based CR, where the problems will be much more challenging.

1.2.2.2 Energy- and Spectral-Efficient Design of OFDMA-based Two-Way Relay

Two-way relay can potentially increase the overall data rate in half-duplex systems compared with one-way relay [15, 16, 17, 18]. Achievable rate regions for two-way *amplify-and-forward* (AF) and *decode-and-forward* (DF) relay channels are derived in [15]. Power allocation at the relay to maximize the weighted sum rate of multiple terminal pairs for two-way relay using frequency and time division multiple access is studied [16]. In [17], relay selection and power allocation for two-way AF relay to minimize the overall transmit power of a terminal pair and multiple relays is investigated. In [18], one hybrid one- and two-way relay transmission scheme is developed, which optimizes the number of bits and

transmission time to minimize the overall energy consumption. In [53], tone permutation and power allocation are performed to maximize the sum capacity over an OFDM-based two-way AF relay. A two-step power allocation approach for OFDM-based two-way AF relay with overall power constraint rather than individual power constraints for the relay and the terminal pair is investigated in [54]. In [55], a hierarchical protocol is proposed for OFDMA-based one- and two-way AF and DF relay, where joint power and subcarrier allocation is conducted using the Lagrange dual decomposition method. In [56], power and subcarrier allocation with subcarrier pairing at the relays is studied for OFDMA-based two-way AF relay, where multiple user terminals communicate with a common base station via multiple relays. In [57], the numbers of active subcarriers and the bits transmitted over each subcarrier at the terminal pair for OFDM-based two-way AF relay, to minimize the overall power consumption. However, limited work considers energy-efficient resource allocation, which directly optimizes the EE utility rather than minimizing the overall power consumption for fixed data rates, for OFDMA-based two-way relay. Moreover, when active subchannel selection is considered for maximizing EE, the underlying EE optimality condition for EE-oriented sequential subchannel assignment has not been discovered or discussed in literature.

1.3 Our Approaches and Thesis Outline

The major goal of this research is to investigate fundamental limits and principles of energy-efficient design and develop novel schemes to improve the EE of wireless communications networks. The proposed energy-efficient schemes will be utilized in various application scenarios, such as resource allocation in OFDMA networks and OFDMA-based CR and two-way relay networks. With these schemes, the EE of wireless communications networks will be significantly enhanced.

As the first step, we investigate the fundamental interrelationship between EE and SE in downlink OFDMA networks. In Chapter 2, we first set up a general EE-SE tradeoff

framework, where the overall EE, SE and per-user minimum rate requirement are all considered. We prove that under this framework, EE is quasiconcave in SE and decreases with SE when SE is large enough. In addition, we discuss some basic properties, such as the impact of channel power gain and circuit power on the EE-SE relation. We also find a tight upper bound and a tight lower bound on the EE-SE curve for general scenarios, which reflect the actual EE-SE relation. We then focus on a special case that priority and fairness are considered and suggest an alternative upper bound, which is proved to be achievable for flat fading channels. We also develop a low-complexity but near-optimal resource allocation algorithm for practical application of the EE-SE tradeoff. The theoretical findings are verified by numeral results. The proposed resource allocation scheme can achieve a flexible and desirable tradeoff between EE and SE.

Energy-efficient resource allocation can save lots of energy of radio access networks in the downlink as well as prolong the battery lifetime and enhance the user experience in the uplink. In Chapter 3, we study the energy-efficient resource allocation in both downlink and uplink cellular networks with OFDMA. For the downlink transmission, the generalized EE is maximized while for the uplink case the minimum individual EE is maximized, both under certain prescribed per-user minimum data rate requirement. For both transmission scenarios, we first provide the optimal solution and then develop a suboptimal but low-complexity approach by exploring the inherent structure and property of the energy-efficient design. For the downlink case, by modifying the original problem, we also find a computationally efficient and numerically tractable upper bound on the EE, which indicates the performance limit and is demonstrated to be quite tight if the number of subcarriers is larger than that of users and motivates us to find a near-optimal approach relying on the quasiconcave relation between the modified EE and transmit power. The proposed energy-efficient design greatly improves EE compared with the conventional spectral-efficient design and the low-complexity suboptimal approaches can achieve a promising tradeoff between performance and complexity.

Although the schemes proposed in Chapter 3 can deal with real-time traffic, it is not appropriate for elastic traffic with average delay requirement. Delay provisioning for delay-sensitive high-data-rate services is an important and challenging issue for future wireless communication networks. In Chapter 4, we study energy-efficient design in downlink OFDMA networks with effective capacity-based delay provisioning for delay-sensitive traffic. By integrating information theory with the concept of effective capacity, we formulate an EE optimization problem with statistical delay provisioning, which is a complicated nonconvex combinatorial fractional programming problem. To solve the problem, we first relax it with an upper bound on the original one and then prove and exploit the quasiconcave property of the EE-versus-transmit power curve, which facilitates the optimal algorithm development. Then, we demonstrate that the resultant solution is quite close to the true optimal value when the number of subcarriers is larger than that of the users. We also analyze the tradeoff between EE and delay, the relationship between spectral-efficient and energy-efficient designs, and the impact of system parameters, including circuit power and delay exponents, on the overall performance. The proposed energy-efficient design scheme greatly improves EE while maintaining the delay requirement.

CR has emerged as one promising technique to achieve high SE. However, EE of CR is not carefully considered but is important. In Chapter 5, we study energy-efficient opportunistic spectrum access strategies for an OFDMA-based CR network with multiple SUs. Both worst EE and average EE are considered and optimized for different emphases and application scenarios. Since the original optimization issues belong to nonconvex integer combinatorial fractional program and are essentially NP-hard for an optimal solution, we use continuous and convex relaxation to modify the problems for somewhat better mathematical tractability. For the relaxed worst-EE-based spectrum access problem, we first demonstrate the joint quasiconcavity of EE on subchannel and power allocation matrices and then develop a framework to find the optimal solution based on efficient root finding

and convex optimization. We also develop a low-complexity alternative for suboptimal solution. The relaxed average-EE-based spectrum access problem is still NP-hard and may have many local optima. We first transform the problem into an equivalent form and introduce a general concave envelope based *branch-and-bound* (B&B) approach to find the global optimal solution. We then exploit the underlying properties of the energy-efficient transmission to speed up the convergence of the B&B approach. Besides, we develop a low-complexity heuristic approach to find a suboptimal solution. The proposed energy-efficient spectrum access strategies significantly boost EE compared with the conventional spectral-efficient spectrum access ones while the low-complexity suboptimal approaches can well balance the performance and complexity.

Two-way relay proves to be more spectral-efficient than direct transmission and one-way relay in many typical scenarios. In Chapter 6, we study energy-efficient resource allocation for OFDMA-based two-way relay, maximizing the aggregated EE utility while provisioning proportional fairness in EE among different terminal pairs. The energy-efficient joint power and subchannel allocation and active subchannel selection problem is mixed-integer combinatorial and further demonstrates to be nonconvex and NP-hard. To approach the performance limit, we first find an upper-bound solution relying on continuous relaxation and the Lagrange dual method. To reduce the computational complexity, we then exploit the hidden concavity and the pseudoconcavity in the subproblems for any fixed subchannel assignment and propose an EE-oriented sequential subchannel assignment policy. Besides, we discover the sufficient condition for early termination of the sequential subchannel assignment without losing the EE optimality. The proposed energy-efficient OFDMA-based two-way relay can achieve much larger EE utility while provisioning proportional fairness in EE among different terminal pairs compared to its spectral-efficient counterpart.

CHAPTER 2

ENERGY- AND SPECTRAL-EFFICIENCY TRADEOFF IN DOWNLINK OFDMA NETWORKS

In this chapter, we address the EE-SE tradeoff issue in downlink OFDMA networks. We build a general EE-SE tradeoff framework, prove that EE is quasiconcave in SE, and discuss the impact of channel power gain and circuit power on the EE-SE relation. We then bound the EE-SE curve for general scenarios by a double-side approximation process, which relies on a tight lower bound and upper bound obtained by *Lagrange dual decomposition* (LDD) and continuous relaxation, respectively. When priority and fairness are considered, we obtain an alternative simple upper bound that is even achievable if all users are with flat-fading channels. To facilitate application of EE-SE tradeoff, we develop a computationally efficient algorithm for resource allocation as well.

The rest of this chapter is organized as follows. In Section 2.1, we describe the system model and establish a general framework for EE-SE tradeoff. In Section 2.2, we first prove the quasiconcavity of the EE-SE curve and analyze the impact of channel power gain and circuit power on the EE-SE relation. We then obtain a tight upper bound and a lower bound on EE-SE curve. We also investigate a specific case that priority and fairness are considered. In Section 2.3, we develop a low-complexity resource allocation algorithm for EE-SE tradeoff. Then, we present numerical results in Section 2.4. Finally, we conclude this chapter in Section 2.5.

2.1 Problem Description

In this section, we introduce the system model and the problem formulation.

2.1.1 System Model

We consider the downlink of a single cell OFDMA network consisting of K active users. The total bandwidth B is equally divided into N subcarriers, each with a bandwidth of

$W = B/N$. Let $\mathcal{K} = \{1, 2, \cdots, K\}$ and $\mathcal{N} = \{1, 2, \cdots, N\}$ denote the sets of active users and all subcarriers, respectively.

In general, one subcarrier can be allocated to multiple users and analysis can be accordingly modeled in this way. For example, in *digital subscriber line* (DSL) broadband access systems, crosstalk (i.e., interference amongst different tones) exists and is thus considered in plenty of literature [58, 59, 60]. In this chapter, we assume that one subcarrier is exclusively assigned to at most one user to avoid interference among different users, the subcarrier frequency spacing is wide enough and *inter-subcarrier interference* (ICI) can be ignored, which coincides with the 3GPP LTE standard and related literature.

Denote the transmit power of user k on subcarrier n as $p_{k,n}$. It is obvious that there is only one $k \in \mathcal{K}$ for each $n \in \mathcal{N}$ such that $p_{k,n} > 0$. Then, the maximum achievable data rate of user k on subcarrier n is

$$r_{k,n} = W \log_2\left(1 + \frac{p_{k,n} g_{k,n}}{N_0 W}\right), \tag{2.1}$$

where $g_{k,n} = |h_{k,n}|^2$ is the channel power gain of user k on subcarrier n, $h_{k,n}$ is the corresponding frequency response and is assumed to be accurately known at the transmitter, and N_0 is the single-sided noise spectral density. Consequently, the overall system throughput and the total transmit power are

$$R = \sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} r_{k,n} \quad \text{and} \quad P = \sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} p_{k,n}, \tag{2.2}$$

respectively, where $\rho_{k,n} \in \{1, 0\}$ indicates whether subcarrier n is allocated to user k (by 1) or not (by 0). Obviously, any feasible subcarrier assignment indictor matrix, $\boldsymbol{\rho} = [\rho_{k,n}]_{K \times N}$, and power allocation matrix, $\boldsymbol{P} = [p_{k,n}]_{K \times N}$, should satisfy

$$\begin{aligned} \boldsymbol{\rho} \in \varrho &\triangleq \left\{ [\rho_{k,n}]_{K \times N} | \sum_{k=1}^{K} \rho_{k,n} \leq 1, \forall\, n \in \mathcal{N}; \rho_{k,n} = \{0, 1\}, \forall\, k \in \mathcal{K}, n \in \mathcal{N} \right\}, \\ \boldsymbol{P} \in \mathcal{P} &\triangleq \left\{ [p_{k,n}]_{K \times N} | p_{k,n} \geq 0, \forall\, k \in \mathcal{K}, \forall\, n \in \mathcal{N}; \sum_{k=1}^{K} \sum_{n=1}^{N} p_{k,n} \leq P_{\max} \right\}, \end{aligned} \tag{2.3}$$

respectively, where $P_{\max}$ represents the peak total transmit power at transmitter.

Circuit technology has made it possible for circuitry of wireless transceivers to consume different power in different operation modes such as sleep, idle, transmit and receive modes [3]. In the transmit/active mode of transmitter, besides transmit power, the energy consumption also includes circuit energy consumption that is incurred by signal processing and active circuit blocks, such as analog-to-digital converter, digital-to-analog converter, synthesizer, and mixer [25, 24, 28, 61]. From [28, 61], circuit energy consumption can be divided into one static (fixed) part and one dynamic part that depends on parameters of active links. Here the transmission associated circuit consumption is modeled as a linear function of throughput [28]

$$P_c = P_s + \xi R, \tag{2.4}$$

where P_s is static circuit power in the transmit mode and ξ is a constant denoting dynamic power consumption per unit data rate. Obviously, the constant circuit consumption model used in [25, 24, 26] is a special case that $\xi = 0$ of the above model. Note that the power loss cost by power amplifier with efficiency $\frac{1}{\zeta}$ to generate the transmit power, P, is not counted in the circuit power, P_c, but will be considered together with transmit power, P, later.

2.1.2 Problem Formulation

For a downlink OFDMA network, EE and SE are defined as

$$\eta_{EE} \triangleq \frac{R}{\zeta P + P_c} \quad \text{and} \quad \eta_{SE} \triangleq \frac{R}{B}, \tag{2.5}$$

respectively, where R, P, and P_c are determined by (2.2) and (2.4), respectively. Throughout this chapter, EE is defined as transmitted bits per unit energy consumption at the transmitter side, where the energy consumption includes transmission energy consumption (ζP times transmission time) and circuit energy consumption (P_c times transmission time) of transmitter in the active mode. Obviously, the larger EE is, the less one transmitted data bit costs.

To obtain a high EE as well as a desirable SE and guarantee *quality-of-service* (QoS) for each user with the limited bandwidth and transmit power resource, it is reasonable

to maximize EE under a satisfying minimum overall throughput requirement, $\check{R}$, and a series of (minimum) rate requirements, $\check{R}_k$'s, which depend on types of the traffic of the corresponding users. Since the capability of providing differentiate services is an important feature for future wireless networks, a heterogeneous traffic model including both real-time and non-real-time traffic is considered in our work. For users with real-time services [62], such as video conferencing and online gaming, fixed rates $\check{R}_i$'s are required. For users with non-real-time services [62], such as file transfers and online video, only minimum rate requirements $\check{R}_j$'s are demanded. Accordingly, the optimization problem can be formulated as

$$\max_{\boldsymbol{\rho}\in\mathcal{Q},\boldsymbol{P}\in\mathcal{P}} \eta_{EE} \left(= \frac{\sum_{k=1}^{K}\sum_{n=1}^{N}\rho_{k,n}r_{k,n}}{\sum_{k=1}^{K}\sum_{n=1}^{N}\rho_{k,n}\left(\zeta p_{k,n}+\xi r_{k,n}\right)+P_s} \right) \tag{2.6a}$$

subject to

$$\sum_{k=1}^{K}\sum_{n=1}^{N}\rho_{k,n}r_{k,n} \geq \check{R}, \tag{2.6b}$$

$$\sum_{n=1}^{N}\rho_{k,n}r_{k,n} = \check{R}_k, \forall\, k \in \mathcal{K}_1, \tag{2.6c}$$

$$\sum_{n=1}^{N}\rho_{k,n}r_{k,n} \geq \check{R}_k, \forall\, k \in \mathcal{K}_2, \tag{2.6d}$$

where $\mathcal{K}_1 = \{1, 2, \cdots, K_0-1\}$ and $\mathcal{K}_2 = \{K_0, K_0+1, \cdots, K\}$ represent the sets of K_0-1 ($K_0 \geq 1$) real-time users and $K - K_0 + 1$ non-real-time users among all the K active users, respectively. Without loss of generality, we can further assume that $\check{R} \geq \sum_{k=1}^{K}\check{R}_k$; Otherwise, (2.6b) can be removed. For specific channel realizations, user and system requirements, it is possible that (2.6) does not have any feasible solution. In this case, transmitter can decrease some users' rate requirements (even to zero) according to their priority and latency requirements, and/or lower the system throughput requirement until (2.6) is feasible. For simplicity, we assume that all constraints in (2.6) can be satisfied simultaneously and $\check{R} \geq \sum_{k=1}^{K}\check{R}_k$ throughout this chapter.

2.2 EE-SE Relation

In this section, we will study the EE-SE relation. We first demonstrate the quasiconcavity of EE in SE, and then discuss the impact of channel power gain and circuit power on the EE-SE relation. We also derive a tight lower bound and a tight upper bound on the EE-SE curve based on LDD and continuous relaxation, respectively. We then focus on the case that priority and fairness are considered.

2.2.1 Fundamentals for EE-SE Relation

The following theorem demonstrates the quasiconcavity of EE, η_{EE}, in the rate vector, $\boldsymbol{R}$, and is proved in Appendix A.1.

Theorem 2.1 *For any given rate vector, $\boldsymbol{R} = [R_k]_{K\times 1}$, achieved with subcarrier allocation matrix, $\boldsymbol{\rho} \in \varrho$, and power allocation matrix, $\boldsymbol{P}$, that satisfy all constraints but not necessarily including the peak transmit power one in (2.6), the maximum EE, $\eta_{EE}^{*}(\boldsymbol{R}) = \max_{\boldsymbol{\rho}\in\varrho, p_{k,n}\geq 0} \eta_{EE}(\boldsymbol{R})$, is strictly quasiconcave in $\boldsymbol{R}$ if there is a sufficiently large number of subcarriers.*

For any continuous and strictly quasiconcave function, there is always a unique global maximum over a finite domain [63, Ch. 8]. Thus, according to Theorem 1, a unique globally optimal EE of (2.6) always exists.

To facilitate practical application of EE-SE tradeoff, we associate EE with SE directly by the next theorem, which is proved in Appendix A.2.

Theorem 2.2 *For any given SE, $\eta_{SE} \geq \frac{\check{R}}{B}$, achieved with subcarrier allocation matrix, $\boldsymbol{\rho} \in \varrho$, and power allocation matrix, $\boldsymbol{P}$, that satisfy all constraints but not necessarily including the peak transmit power one in (2.6), the maximum EE, $\eta_{EE}^{*}(\eta_{SE}) = \max_{\boldsymbol{\rho}\in\varrho, p_{k,n}\geq 0} \eta_{EE}(\eta_{SE})$, is strictly quasiconcave in η_{SE} if there is a sufficiently large number of subcarriers. Moreover, in the SE region $\left[\frac{\check{R}}{B}, \frac{\hat{R}}{B}\right]$, the EE, $\eta_{EE}^{*}(\eta_{SE})$*

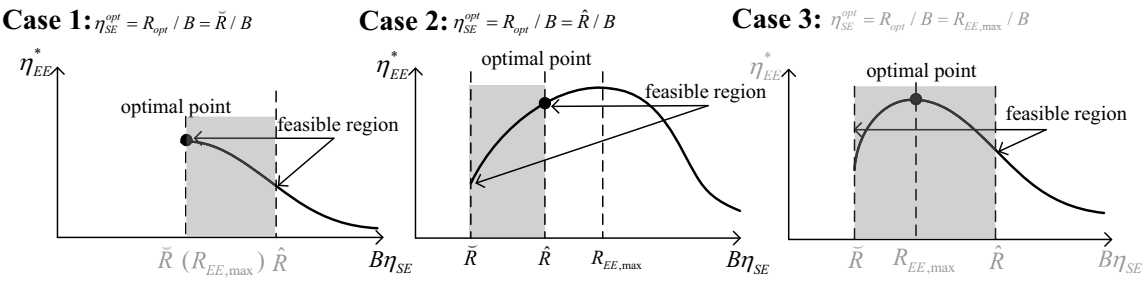


Figure 2.1. EE-SE relation in downlink OFDMA.

(i) strictly decreases with η_{SE} *and is maximized at* $\eta_{SE} = \frac{\check{R}}{B}$ *if*

$$\left.\frac{\mathrm{d}\eta_{EE}^*(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\frac{\check{R}}{B}} \leq 0,$$

(ii) strictly increases with η_{SE} *and is maximized at* $\eta_{SE} = \frac{\hat{R}}{B}$ *if*

$$\left.\frac{\mathrm{d}\eta_{EE}^*(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\frac{\check{R}}{B}} > 0 \quad \text{and} \quad \left.\frac{\mathrm{d}\eta_{EE}^*(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\frac{\hat{R}}{B}} \geq 0,$$

(iii) first strictly increases and then strictly decreases with η_{SE} *and is maximized at* $\eta_{SE} = \frac{R_{EE,max}}{B}$ *if*

$$\left.\frac{\mathrm{d}\eta_{EE}^*(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\frac{\check{R}}{B}} > 0 \quad \text{and} \quad \left.\frac{\mathrm{d}\eta_{EE}^*(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\frac{\hat{R}}{B}} < 0,$$

where $\hat{R}$ *is the maximum throughput under all constraints in (2.6), and* $R_{EE,max}$ *is the throughput that corresponds to the maximum EE,* $\eta_{EE}^{\max}$*, under all constraints but not necessarily including the peak transmit power one in (2.6).*

Theorem 2.2 demonstrates the quasiconcavity of EE on SE and reconfirms the existence and the uniqueness of the globally optimal EE, η_{EE}^{opt}, of (2.6). In contrary to Theorem 2.1, it is from a uniform SE (or total rate) perspective rather than that of a vector of split user rates, which makes it easier to track the EE-SE relation.

Figure 2.1 illustrates three possible EE-SE curves discussed in Theorem 2.2. In Case 1, the optimal EE, η_{EE}^{opt}, is reached when R_{opt} equals $\check{R}$; in Case 2, the optimal EE, η_{EE}^{opt}, is reached when R_{opt} equals $\hat{R}$; in Case 3, the optimal EE, η_{EE}^{opt}, is reached when R_{opt} equals

$R_{EE,max}$, where R_{opt} is the optimal throughput of (2.6). Note that $\eta_{EE}^{\max}$ equals η_{EE}^{opt} in Cases 1 and 3 while it is greater than the latter in Case 2. Furthermore, Theorem 2.2 offers a simple and adaptive way to determine online a desirable and feasible minimum throughput requirement, $\check{R}$, in practical situations. Initially, set $\check{R} = \sum_{k=1}^{K} \check{R}_k$. Then, figure out which one of the three cases occurs. In Case 1, $\check{R}$ should be eventually set in $[\sum_{k=1}^{K} \check{R}_k, \hat{R}]$ and the system will just work at this point once it is determined. In Case 2, $\check{R}$ should be eventually set $\hat{R}$ and the system will just work at this point. In Case 3, $\check{R}$ should be eventually set in $[R_{EE,max}, \hat{R}]$ and the system will just work at this point. More importantly, as a result of this quasiconcavity, problem (2.6) can be decomposed into two layers and solved iteratively,

(i) Inner layer: For a given SE, η_{SE}, find the maximum EE, $\eta_{EE}^{*}(\eta_{SE})$, and its derivative, $\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$.

(ii) Outer layer: Find the optimal EE, η_{EE}^{opt}, by bisection search like the GABS algorithm in [26].

The corresponding *joint inner- and outer-layer optimization* (JIOO) algorithm is listed in Table 2.1. As shown there, the outer-layer procedure is simple and it takes at most $\lceil \log_2 \left(\frac{(\alpha-1)\eta_{SE}^{opt}}{\epsilon} - 1 \right) \rceil$ iterations to converge to an extent that $|\eta_{SE} - \eta_{SE}^{opt}| < \epsilon$, where $\lceil x \rceil$ denotes the smallest integer not less than x [26]. The key of the JIOO algorithm lies in the inner-layer algorithm that finds $\eta_{EE}^{*}(\eta_{SE})$ and $\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$ and will be discussed in detail in the following sections.

The impact of channel power gain and circuit power on the EE-SE relation is summarized in the following statements and proved in Appendix A.3. Note that the $\eta_{EE}^{*}\left(\eta_{SE}\right)$-versus-$\eta_{SE}$ curve is exclusively referred to as the EE-SE curve throughout this chapter.

Property 2.1 *For any fixed SE, the EE is non-decreasing with channel power gain. Consequently, both the optimal EE, η_{EE}^{opt}, and the global maximum EE, $\eta_{EE}^{\max}$, are non-decreasing with channel power gain $g_{k,n}$.*

Table 2.1. Joint Inner- and Outer-layer Optimization (JIOO) Algorithm.

Algorithm *JIOO*
Input: initial value of SE: $\eta_{SE}^{(0)} = \frac{\check{R}}{B}$
Output: optimal subcarrier and power allocation matrices, $\boldsymbol{\rho}_{opt}$ and $\boldsymbol{P}_{opt}$

1. $\eta_{SE}^{(1)} = \eta_{SE}^{(0)}$, $d_1 \leftarrow \left.\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{(1)}}$
2. **if** $d_1 \leq 0$
3. **then** ($*$ Case 1 as shown in Figure 2.1 $*$)
4. **return** current $\boldsymbol{\rho}$ and $\boldsymbol{P}$ as $\boldsymbol{\rho}_{opt}$, $\boldsymbol{P}_{opt}$, respectively
5. **else** ($*$ Cases 2 and 3 as shown in Figure 2.1 $*$)
6. $\eta_{SE}^{(2)} \leftarrow \eta_{SE}^{(1)}$, $\eta_{SE}^{(1)} \leftarrow \kappa\eta_{SE}^{(1)}$, $P_1 \leftarrow P^{*}(\eta_{SE}^{(1)})$, $d_1 \leftarrow \left.\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{(1)}}$, where $P^{*}(\eta_{SE}^{(1)}) = \frac{B\eta_{SE}^{(1)}}{\zeta\eta_{EE}^{*}\left(\eta_{SE}^{(1)}\right)} - \frac{P_c}{\zeta}$ and $\kappa > 1$, e.g., $\kappa = 2$
7. **while** $d_1 > 0$ && $P_1 < P_{\max}$
8. **do** $\eta_{SE}^{(2)} \leftarrow \eta_{SE}^{(1)}$, $\eta_{SE}^{(1)} \leftarrow \kappa\eta_{SE}^{(1)}$, $P_1 \leftarrow P^{*}(\eta_{SE}^{(1)})$, $d_1 \leftarrow \left.\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{(1)}}$
9. **while** no convergence ($*$ bisection search $*$)
10. **do** $\eta_{SE}^{opt} \leftarrow \frac{\eta_{SE}^{(1)}+\eta_{SE}^{(2)}}{2}$, $P_{opt} \leftarrow P^{*}(\eta_{SE}^{(opt)})$, $d_{opt} \leftarrow \left.\frac{\mathrm{d}\eta_{EE}^{*}\left(\eta_{SE}^{opt}\right)}{\mathrm{d}\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{(1)}}$, where $P^{*}(\eta_{SE}^{(opt)}) = \frac{B\eta_{SE}^{(opt)}}{\zeta\eta_{EE}^{*}\left(\eta_{SE}^{(opt)}\right)} - \frac{P_c}{\zeta}$
11. **if** $d_{opt} > 0$ && $P_{opt} < P_{\max}$
12. **then** $\eta_{SE}^{(2)} \leftarrow \eta_{SE}^{opt}$
13. **else** $\eta_{SE}^{(1)} \leftarrow \eta_{SE}^{opt}$
14. **return** current $\boldsymbol{\rho}$ and $\boldsymbol{P}$ as $\boldsymbol{\rho}_{opt}$, $\boldsymbol{P}_{opt}$, respectively

Property 2.2 *For any fixed SE, the EE strictly decreases with circuit power. Consequently, both the optimal EE, η_{EE}^{opt}, and the global maximum EE, $\eta_{EE}^{\max}$, strictly decrease with the circuit power, P_c. Besides, the optimal SE, $\eta_{SE}^{opt} \overset{\text{def}}{=} \frac{R_{opt}}{B}$, is non-decreasing with the static circuit power, P_s, while $\eta_{SE}^{\max} \overset{\text{def}}{=} \frac{R_{EE,max}}{B}$ strictly increases with the static circuit power, P_s. However, both η_{SE}^{opt} and $\eta_{SE}^{\max}$ are independent of the dynamic circuit power factor ξ.*

Property 2.1 coincides with intuition. From it, the scheduling of users with better channel quality, such as users near base stations, helps to improve the system EE as well as SE. Property 2.2 is sort of unexpected and not that straightforward as for the independence of $\eta_{SE}^{\max}$ and η_{SE}^{opt} on the dynamic circuit power factor ξ. From it, high SE may be utilized without much loss in EE in the case that static circuit power and/or dynamic circuit power factor are large. This is because large static circuit power indicates large $\eta_{SE}^{\max}$ and η_{SE}^{opt} while

large dynamic power factor makes the EE-SE curve be low and flat (EE is insensitive to the change of SE) although it does not impact the locations of $\eta_{SE}^{\max}$ and η_{SE}^{opt}.

Property 2.3 *The optimal EE, η_{EE}^{opt}, is not necessarily achieved at $\eta_{SE} = \frac{\check{R}}{B}$ even if circuit power, P_c, is zero.*

We prove it by a counterexample. Let us consider a virtual downlink OFDMA system with parameters $B = 10$, $N = 10$, $K = 2$, $\sigma^2 = 1$, $\frac{g_{1,n}}{N_0} = \frac{g_1}{N_0} = 1$ and $\frac{g_{2,n}}{N_0} = \frac{g_2}{N_0} = 10$ for $n = 1, 2, \cdots, N$, $\check{R}_1 = 10$, $\check{R}_2 = 0$, $\check{R} = 10$, $P_c = 0$ and $P_{\max} = 20$.

In the case that $R = \check{R}$, it is obvious that the optimal scheme is to allocate all subcarriers to user 1 and distribute equally the power on each subcarrier. The EE is accordingly

$$\eta_{EE}^{*}\left(\check{R}\right) = \frac{R}{N\frac{N_0 W}{g_1}2^{\left(\frac{R}{N}-1\right)}} = 1.$$

In the case that $R = 20$, it is easy to verify that the optimal scheme is to allocate $m_1 = 7$ subcarriers to user 1 to realize $R_1 = \check{R}_1 = 10$ while allocating $m_2 = 3$ subcarriers to user 2 to realize $R_2 = R - \check{R}_1 = 10$. The EE is

$$\eta_{EE}^{*}(20) = \frac{R}{\sum_{k=1}^{K} m_k \frac{N_0 W}{g_k} 2^{\left(\frac{R_k}{m_k}-1\right)}} = 1.373.$$

Therefore, we have $\eta_{EE}^{*}\left(\check{R}\right) < \eta_{EE}^{*}(20) \leq \eta_{EE}^{opt}$, which supports Property 2.3.

Property 2.3 indicates that, in the case that transmit power (absolutely) dominates power consumption, transmitting with minimum power to merely realize the minimum throughput requirement and guarantee the QoS of each user is not necessarily the most energy-efficient scheme. In fact, it is easy to verify that both Theorems 2.1 and 2.2 still hold without considering the circuit power in (2.6). And it is well worth noting that in the single user case ($K = 1$), η_{EE}^{opt} is always achieved at the minimum throughput, which is consistent with the conclusion in [26] and can be proved similarly.

Note that all conclusions (Theorems and Properties) in Section 2.2.1 expect the independence of η_{SE}^{opt} and $\eta_{SE}^{\max}$ on dynamic circuit power factor ξ also hold in a more general

case that circuit power is a convex function of throughput, which can be similarly proved following the linear ones.

2.2.2 Bounds on the EE-SE Curve

In this part, upper and lower bounds on the EE-SE curve are studied. As indicated before, the solution of (2.6) now relies on finding $\eta_{EE}^{*}(\eta_{SE})$ and $\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$. Since the exact solution is too complicated to obtain in reality, we will use LDD to approximately solve it, which has been used in [64, 65, 66] for similar problems and demonstrated to be quite accurate with affordable computational complexity. For any given throughput, $R \geq \check{R}$, the Lagrange dual problem of minimizing the total transmit power can be expressed as

$$\max_{\substack{\boldsymbol{\lambda}_1 \geq \mathbf{0}, \boldsymbol{\lambda}_2 \geq \mathbf{0}, \\ \lambda_3 \geq 0}} \min_{\boldsymbol{\rho} \in \mathcal{Q}, \boldsymbol{P} \in \mathcal{P}} \left\{ \sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} p_{k,n} + \sum_{n=1}^{N} \lambda_{1,n} \left(\sum_{k=1}^{K} \rho_{k,n} - 1 \right) + \sum_{k=1}^{K} \lambda_{2,k} \left(\check{R}_k - \sum_{n=1}^{N} \rho_{k,n} r_{k,n} \right) + \lambda_3 \left(R - \sum_{k=1}^{K_0 - 1} \check{R}_k - \sum_{k=K_0}^{K} \sum_{n=1}^{N} \rho_{k,n} r_{k,n} \right) \right\}, \tag{2.7}$$

where $\lambda_{1,n}$, $\lambda_{2,k}$, and λ_3 are the introduced Lagrange multipliers, and $\boldsymbol{\lambda}_1 = [\lambda_{1,1}, \lambda_{1,2}, \cdots, \lambda_{1,N}]^T$ and $\boldsymbol{\lambda}_2 = [\lambda_{2,1}, \lambda_{2,2}, \cdots, \lambda_{2,K}]^T$, respectively. And $\geq$ is used as a component-wise inequality. Note that we drop the peak power constraint in (2.6) here since it will be imposed by the outer-layer processing as shown in the JIOO algorithm.

Problem (2.7) can be decomposed into two layers [64, 65, 66, 67]. In the lower layer, the subordinate problem for subcarrier $n \in \mathcal{N}$ is

$$U_n = \min_{\boldsymbol{\rho} \in \mathcal{Q}, \boldsymbol{P} \in \mathcal{P}} \sum_{k=1}^{K} \rho_{k,n} u_{k,n}, \tag{2.8a}$$

subject to

$$\begin{aligned} u_{k,n} &= p_{k,n} + \lambda_{1,n} - \lambda_{2,k} r_{k,n}, \text{ if } k \in \mathcal{K}_1, \\ u_{k,n} &= p_{k,n} + \lambda_{1,n} - (\lambda_{2,k} + \lambda_3) r_{k,n}, \text{ if } k \in \mathcal{K}_2. \end{aligned} \tag{2.8b}$$

The master problem in the upper layer will be

$$\max_{\substack{\boldsymbol{\lambda}_1 \geq \mathbf{0}, \boldsymbol{\lambda}_2 \geq \mathbf{0}, \\ \lambda_3 \geq 0}} \left\{ \sum_{n=1}^{N} U_n - \sum_{n=1}^{N} \lambda_{1,n} + \sum_{k=1}^{K} \lambda_{2,k} \check{R}_k + \lambda_3 \left(R - \sum_{k=1}^{K_0 - 1} \check{R}_k \right) \right\}. \tag{2.9}$$

Problems (2.8) and (2.9) can be solved by iteratively exchanging information between them and updating the Lagrange multipliers. In the lower layer, if $\rho_{k,n} = 1$, the transmit power needed by water-filling is

$$\begin{aligned} p_{k,n} &= W\left[\frac{\lambda_{2,k}}{\ln 2} - \frac{N_0}{g_{k,n}}\right]^+, \text{ if } k \in \mathcal{K}_1, \\ p_{k,n} &= W\left[\frac{(\lambda_{2,k} + \lambda_3)}{\ln 2} - \frac{N_0}{g_{k,n}}\right]^+, \text{ if } k \in \mathcal{K}_2, \end{aligned} \tag{2.10}$$

where $[x]^+ = \max(x, 0)$. Once the above equation is solved for all $k \in \mathcal{K}$, $p_{k,n}$'s are substituted back into (2.8b), and $\rho_{k,n}$'s are set to 1 only for the user with minimum $u_{k,n}$ while 0 for all the other users. The master problem (2.9) is in charge of determining the Lagrange multipliers and feeding them back to the N subproblems in the lower layer. Note that $\boldsymbol{\lambda}_1$ is actually useless in our optimization, which can be initialized any value and remains unchanged. The other Lagrange multipliers can be updated using the ellipsoid method [58] or subgradient method [64, 58, 65, 66, 67] as

$$\lambda_{2,k}^{(i+1)} = \left[\lambda_{2,k}^{(i)} - s^{(i)}\left(\sum_{n=1}^{N} \rho_{k,n}^{(i)} r_{k,n}^{(i)} - \check{R}_k\right)\right]^+, \tag{2.11a}$$

$$\lambda_3^{(i+1)} = \left[\lambda_3^{(i)} - s^{(i)}\left(\sum_{k=K_0}^{K}\sum_{n=1}^{N} \rho_{k,n}^{(i)} r_{k,n}^{(i)} - \left(R - \sum_{k=1}^{K_0-1} \check{R}_k\right)\right)\right]^+ \tag{2.11b}$$

until convergence, where $s^{(i)}$ is a sufficiently small positive stepsize for the ith iteration. For example, $s^{(i)}$ can be determined in a diminishing way, such as $s^{(i)} = \frac{0.1}{\sqrt{i}}$ [58, 66]. More powerful Lagrange multiplier updating methods can be applied by smoothing the objective function and using the optimal gradient with further optimized stepsizes, which may lead to even faster convergence [60]. The topic on advanced Lagrange multiplier updating techniques is beyond the scope of this work, refer to [60] and references therein for details. Figure 2.2 sketches the procedure of the LDD algorithm.

Since the primal problem of (2.7) is nonconvex, there can be many local optima and there is possibly a duality gap when the problem is solved by LDD. Fortunately, it is proved that strong duality holds and the optimal duality gap goes to zero as the number of subcarriers approaches infinity [58, 59]. And the duality gap is already quite small for practical

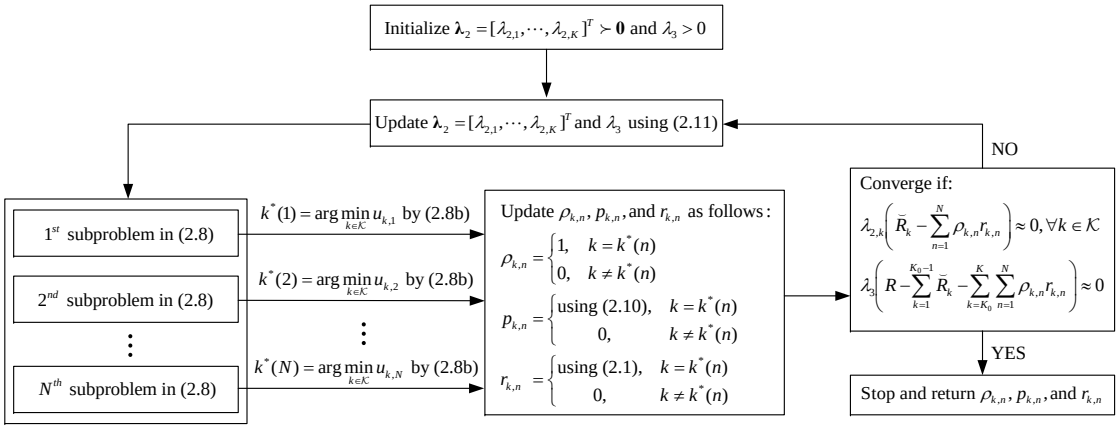


Figure 2.2. Structure of the LDD approach.

number of subcarriers as observed by simulation [58, 64, 65, 66, 67, 59, 68]. For example, in a downlink OFDMA network with 2 users and 33 used data subcarriers, an average duality gap of less than 10^{-5} can be achieved [68]. Nevertheless, the duality gap may exist and be large for particular channel realizations, user and system requirements. To guarantee the solution feasible, we further find a tight and achievable upper bound on the minimum transmit power based on the subcarrier allocation strategy derived from the LDD, which is stated in Theorem 2.1 and proved in Appendix A.4.

Theorem 2.3 *For any fixed throughput, R, and any corresponding feasible subcarrier allocation strategy, $\boldsymbol{\rho} = [\rho_{k,n}]_{K,N}$, the minimum transmit power can be achieved when power is distributed in a two-stage-water-filling fashion. In the first stage, power is distributed individually among the subcarriers of each user, $\mathcal{S}_k$, by water-filling, to merely fulfill its own (minimum) rate requirement, $\check{R}_k$; in the second stage, extra power is then distributed among all subcarriers of the non-real-time users (each has an initial water level due to the first phase processing) by water-filling till the throughput, R, is achieved.*

According to Theorem 2.3, we can directly conclude that for the subcarrier allocation strategy derived from the dual decomposition, the transmit power obtained by Theorem 2.3 is an achievable upper bound on the minimum transmit power in the primal problem of (2.7).

One the other hand, a lower bound on the minimum transmit power in the primal problem of (2.7) can be obtained by relaxing the channel allocation indictors $\rho_{k,n}$ into continuous real variables over [0, 1] indicating the portion of subcarrier n used by user k and modifying (2.1) into [69, 70, 71]

$$r_{k,n} = W \log_2\left(1 + \frac{p_{k,n} g_{k,n}}{\rho_{k,n} W N_0}\right). \tag{2.12}$$

With these manipulations, the primal problem of the modified problem (2.7) is converted into a convex minimization problem over a convex set, thus, it can be optimally solved by standard convex optimization techniques such as interior point method. In the optimal solution, only a small number of subcarriers are shared among users as $\rho_{k,n}$ is mostly either 1 or 0 for $K \ll N$ [70, 71]. Thus, this lower bound on total transmit power is tight when the number of subcarriers is large enough.

Using the definitions $\eta_{EE} = \frac{R}{P+P_c}$ and $\eta_{SE} = \frac{R}{B}$, the upper bound and the lower bound on the total transmit power, P, correspond to a lower bound and an upper bound on EE, $\eta^*_{EE}(\eta_{SE})$, respectively. Now we still need $\frac{\mathrm{d}\eta^*_{EE}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$ to implement the JIOO algorithm. The next theorem provides a simple way to determine the sign of $\frac{\mathrm{d}\eta^*_{EE}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$ by exploiting the intrinsic relationship between $\frac{\mathrm{d}\eta^*_{EE}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$ and $\frac{\mathrm{d}P^*(R)}{\mathrm{d}R}$, which is proved in Appendix A.5.

Theorem 2.4 *For the maximum EE, $\eta^*_{EE}(\eta_{SE})$, defined in Theorem 2.2, its first derivative, $\frac{\mathrm{d}\eta^*_{EE}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$, satisfies*

$$\frac{\mathrm{d}\eta^*_{EE}(\eta_{SE})}{\mathrm{d}\eta_{SE}} \begin{cases} > 0 & \text{if } \eta^*_{EE}(\eta_{SE}) < \frac{1}{\zeta \frac{\mathrm{d}P^*(R)}{\mathrm{d}R} + \xi}, \\ = 0 & \text{if } \eta^*_{EE}(\eta_{SE}) = \frac{1}{\zeta \frac{\mathrm{d}P^*(R)}{\mathrm{d}R} + \xi}, \\ < 0 & \text{if } \eta^*_{EE}(\eta_{SE}) > \frac{1}{\zeta \frac{\mathrm{d}P^*(R)}{\mathrm{d}R} + \xi}, \end{cases} \tag{2.13}$$

where $P^*(R) \overset{\text{def}}{=} \frac{R}{\zeta \eta^*_{EE}(R)} - \frac{P_s + \xi R}{\zeta}$ is the minimum transmit power required for throughput $R(= B\eta_{SE})$. Moreover, $\frac{\mathrm{d}P^*(R)}{\mathrm{d}R} \equiv \min_{k \in \mathcal{K}_2, n \in \mathcal{S}_k} \frac{\ln 2}{W}\left(p^*_{k,n} + \frac{N_0 W}{g_{k,n}}\right)$, where $p^*_{k,n}$ represents the transmit power of user k on subcarrier n corresponding to $P^*(R)$.

Interestingly, although we can hardly derive a closed-form expression of the derivative function, $\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$, we can easily and accurately determine its sign given the solution of the previous transmit power minimization problem. Theorem 2.4 has clear physical insights and coincides with intuition. That is, for $\eta_{EE}^{*}(\eta_{SE}) = \frac{R}{\zeta P^{*}(R)+\xi R+P_s}$ with a structure of $\frac{\texttt{numerator}}{\texttt{denominator}}$, its first derivative depends on the relative size of $\frac{R}{\zeta P^{*}(R)+\xi R+P_s}$ and $\frac{\Delta R}{\zeta \Delta P^{*}(R)+\xi \Delta R}$, where ΔR and $\Delta P^{*}(R)$ are infinitely small increments of throughput and corresponding transmit power, respectively. For example, if $\frac{R}{\zeta P^{*}(R)+\xi R+P_s}$ is smaller than $\frac{\Delta R}{\zeta \Delta P^{*}(R)+\xi \Delta R}$, then the derivative $\frac{\mathrm{d}\eta_{EE}^{*}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$ is positive. This result is in accord with the intuition that if the additional total power cost of increasing certain amount of throughput is small enough, increasing throughput will enhance EE at the same time. In addition, $\frac{\mathrm{d}P^{*}(R)}{\mathrm{d}R}$ at every specific $R(\geq \check{R})$ depends only on the non-real-time user with the minimum water-level (this indicates it has the maximum derivative) but not on the other non-real-time users and real-time users. The strict rate constraints for real-time users make them irrelevant to $\frac{\mathrm{d}P^{*}(R)}{\mathrm{d}R}$ at every specific R although some of them might have better channel quality and thus lead to the same amount of throughput improvement at a even lower additional transmit power cost than the best non-real-time user does.

2.2.3 Priority and Fairness Issues

For many practical scenarios, ability to provide different service priority and fairness among users is important. Therefore, in this part, we incorporate these issues into (2.6).

Let $\tilde{R} = R - \sum_{k=1}^{K} \check{R}_k$, then the two rate constraints in (2.6) can be rewritten as

$$\sum_{n=1}^{N} \rho_{k,n} r_{k,n} = \check{R}_k + \omega_k \tilde{R},\ \forall k \in \mathcal{K}, \tag{2.14a}$$

$$\omega_k = 0,\ \forall\, k \in \mathcal{K}_1, \tag{2.14b}$$

$$\sum_{k=K_0}^{K} \omega_k = 1,\ \omega_k \geq 0,\ \forall\, k \in \mathcal{K}_2, \tag{2.14c}$$

where ω_k is the weight factor for user k. We can determine the weights according to user traffic types, fairness and priority requirements. For users with real-time services, ω_i can

be simply set to zero. For users with non-real time services, we can prioritize them and enforce certain notations of fairness by adjusting ω_j. For example, we can determine ω_j by making user rates proportional to a set of predetermined values among non-real-time users [72].

No matter what principle of fairness or priority is defined, with correspondingly predetermined weight vector, $\boldsymbol{\omega}$, the equivalent problem of (2.6) given the throughput R can be regarded as the conventional *margin adaptation* (MA) optimization problem [64, 69, 73, 74] as follows [1]

$$\min_{\boldsymbol{\rho}\in\mathcal{Q},\boldsymbol{P}\in\mathcal{P}} \sum_{k=1}^{K}\sum_{n=1}^{N}\rho_{k,n}p_{k,n} \tag{2.15a}$$

subject to

$$\sum_{n=1}^{N}\rho_{k,n}r_{k,n} = \check{R}_k + \omega_k\tilde{R}, \forall\, k \in \mathcal{K}, \tag{2.15b}$$

Compared with (2.7), this problem is much less complicated because it does not need to distribute extra rate among different users. Although LDD can be still employed as before, we can easily derive another upper bound to roughly estimate the EE-SE relation in this case. Motivated by the BABS algorithm [73] for finding distribution of subcarriers in flat-fading channels, we suggest the following transmit-power-lower-bound (T-PLB) algorithm. The basic idea is to relax problem (2.15) by eliminating the constraints $\sum_{k=1}^{K}\rho_{k,n} \le 1, \forall\, n \in \mathcal{N}$, and iteratively assigning the subcarriers until the summation of the number of assigned subcarriers of all the users, $\sum_{k=1}^{K} m_k$, equals N. In initialization, each user is assigned its optimal subcarrier. Then, in each iteration, each user finds its optimal subcarrier among its own unobtained ones and calculates the transmit-power saving with the additional subcarrier. Only the user with the most transmit-power saving will be assigned its currently favorite subcarrier in this iteration. Note that some subcarriers may be assigned to no user while some may be assigned to multiple users. Thus, the result is likely

[1] Another popular category of resource allocation problems in OFDMA networks is *rate adaptation* (RA) that maximizes overall throughput under per-user QoS requirement and system resource constraints [36].

Table 2.2. Transmit-Power-Lower-Bound (TPLB) Algorithm.

1. Initialize: $\mathcal{S}_k \leftarrow \varnothing$, $m_k \leftarrow 0$, $\mathcal{N}_k \leftarrow \mathcal{N}$, $\forall k \in \mathcal{K}$.
2. $\hat{n}_k \leftarrow \arg\max_{n\in\mathcal{N}} g_{k,n}$, $P_k \leftarrow f(R_k, \{\hat{n}_k\})$, $\forall\, k \in \mathcal{K}$.
   Assign and update: $\mathcal{S}_k \leftarrow \{\hat{n}_k\}$, $\mathcal{N}_k \leftarrow \mathcal{N}_k \setminus \{\hat{n}_k\}$, $m_k \leftarrow m_k + 1$, $\forall k \in \mathcal{K}$.
3. $\hat{n}_k \leftarrow \arg\max_{n\in\mathcal{N}_k} g_{k,n}$, $\Delta P_k \leftarrow P_k - f(R_k, \mathcal{S}_k \cup \{\hat{n}_k\})$, $\forall\, k \in \mathcal{K}$.
   $\hat{k} \leftarrow \arg\max_{k\in\mathcal{K}} \Delta P_k$.
   Assign and update: $\mathcal{S}_{\hat{k}} \leftarrow \mathcal{S}_{\hat{k}} \cup \{\hat{n}_{\hat{k}}\}$, $\mathcal{N}_{\hat{k}} \leftarrow \mathcal{N}_{\hat{k}} \setminus \{\hat{n}_{\hat{k}}\}$, $m_{\hat{k}} \leftarrow m_{\hat{k}} + 1$, $P_{\hat{k}} \leftarrow P_{\hat{k}} - \Delta P_{\hat{k}}$.
4. Repeat Step 3 until $\sum_{k=1}^{K} m_k = N$.

not feasible or achievable. The procedure of the TPLB algorithm is summarized in Table 2.2 and its property is given in Theorem 2.5 and proved in Appendix A.6.

Theorem 2.5 *The transmit power obtained by the TPLB algorithm is indeed a lower bound on the optimal solution to (2.15). Furthermore, it is achievable if all users are with flat-fading channels.*

According to Theorem 2.5, the more correlation among the subcarriers of each user is, the tighter the lower bound is. Besides, this algorithm only needs $O(NK)$ times of water-filling to complete while the LDD method needs at most $O(NK/(1/\varepsilon^2))$ times of water-filling to converge to ε-optimality [60, 67].

On the other hand, the above lower bound on the total transmit power, P, corresponds to an upper bound on EE, $\eta^*_{EE}(\eta_{SE})$. And in this case, to obtain the sign of $\frac{\mathrm{d}\eta^*_{EE}(\eta_{SE})}{\mathrm{d}\eta_{SE}}$ to implement the JIOO algorithm, we can follow a similar way like the one for the general case indicated in Theorem 2.4. That is, use (2.13) with a modification that $\frac{\mathrm{d}P^*(R)}{\mathrm{d}R} = \sum_{k\in\mathcal{K}_2} \frac{\omega_k \ln 2}{W} \min_{n\in\mathcal{S}_k}\left(p^*_{k,n} + \frac{N_0 W}{g_{k,n}}\right)$.

2.3 Low-complexity Algorithm Design

The LDD can provide a quite accurate solution to (2.15). However, it needs $O(NK/(1/\varepsilon^2))$ times of water-filling to converge in the worst case and its complexity is still very high.

Table 2.3. Maximum-Power-Decrease-First (MPDF) Algorithm.

1. Initialize: $\mathcal{K}_E \leftarrow \mathcal{K}$; $\mathcal{S}_k \leftarrow \varnothing, m_k \leftarrow 0, \forall\, k \in \mathcal{K}$.
2. Calculate benchmarks: $\check{n}_k \leftarrow \arg\min_{n\in\mathcal{N}} g_{k,n}$, $P_k \leftarrow f(R_k, \check{n}_k)$, $\forall\, k \in \mathcal{K}$.
3. $\hat{n}_k \leftarrow \arg\max_{n\in\mathcal{N}} g_{k,n}$, $\Delta P_k \leftarrow P_k - f(R_k, \mathcal{S}_k \cup \{\hat{n}_k\})$, $\forall\, k \in \mathcal{K}$. $\hat{k} \leftarrow \arg\max_{k\in\mathcal{K}} \left(1 + \delta\Big(N - \sum_{k=1}^{K} m_k - |\mathcal{K}_E|\Big)\delta(m_k)\right)\Delta P_k$.
   Assign and update: $\mathcal{S}_{\hat{k}} \leftarrow \mathcal{S}_{\hat{k}} \cup \{\hat{n}_{\hat{k}}\}$, $\mathcal{N} \leftarrow \mathcal{N} \setminus \{\hat{n}_{\hat{k}}\}$, $m_{\hat{k}} \leftarrow m_{\hat{k}} + 1$, $P_{\hat{k}} \leftarrow P_{\hat{k}} - \Delta P_{\hat{k}}$, $\mathcal{K}_E \leftarrow \mathcal{K}_E \setminus \{\hat{k}\}$.
4. Repeat Step 3 until $\sum_{k=1}^{K} m_k = N$ or $\max_{k\in\mathcal{K}} \Delta P_k = 0$.

Therefore, we investigate a low-complexity but near-optimal algorithm to allocate resource for EE-SE tradeoff for (2.15) and also for (2.6).

The key idea of our greedy approach is to modify the TPLB algorithm and make it feasible. Different from the TPLB algorithm that assigns each user a subcarrier at the beginning to ensure the nonempty of its assigned subcarrier set at last, in initialization of this approach, each user is only virtually assigned its worst subcarrier, and its transmit power needed in this situation will be used as a benchmark to measure how much power can be saved if the user is actually assigned a subcarrier when he has none. Then, in each iteration, each user finds its optimal subcarrier among the unassigned ones (not among its own unobtained ones like in TPLB) and calculates its power decrease with the additional subcarrier. Only the user with the maximum power decrease (transmit-power saving) will be assigned its currently favorite subcarrier in this iteration. The above iteration process will proceed until all the subcarriers have been assigned or no user needs additional subcarriers. Note that to guarantee each user is assigned at least one subcarrier at last, when the the number of unassigned subcarriers equals the number of users without any subcarriers, the subcarrier assignment should be implemented among such empty users. This algorithm is referred as *maximum-power-decrease-first* (MPDF) and is sketched in Table 2.3.

In Step 3, for such users that their optimal subcarriers in the last iteration are still available (has not been allocated), there is no need to search for their optimal subcarriers and

calculate ΔP_k's this time. This could potentially further reduce computational complexity.

Next, we give a theorem to grasp insight into the MPDF algorithm.

Theorem 2.6 *The power required by water-filling with a fixed target rate is convex and non-increasing with the number of subcarriers assigned if the subcarriers are added in descending order of the power gains. (The proof is omitted.)*

The MPDF algorithm operates in a user-centric fashion, where each user considers its own optimal available subcarrier each time, and makes assignment decision directly towards minimizing the transmit power. Meanwhile, from Theorem 2.6, the MPDF algorithm naturally prevents one user getting too many subcarriers since benefit of acquiring subcarriers is decreasing. Note that the MPDF algorithm is equivalent to the TPLB algorithm for flat-fading channels and achieves the optimal result in that case.

Another resource allocation algorithm, i.e., the DPRA algorithm, developed in [74], works in a subcarrier-centric fashion. It initializes each user's subcarrier set with all the N subcarriers and then eliminates one subcarrier from all but one user each time until each subcarrier remains for only one user. Compared with the DPRA algorithm, our MPDF algorithm better reflects the impact of good subcarriers by allowing each user to choose his own optimal one. And more importantly, our algorithm performs water-filling with a small number of subcarriers and needs no iteration to obtain the water level due to the decreasing order of channel gain. Both algorithms need $O(NK)$ times of water-filling in total.

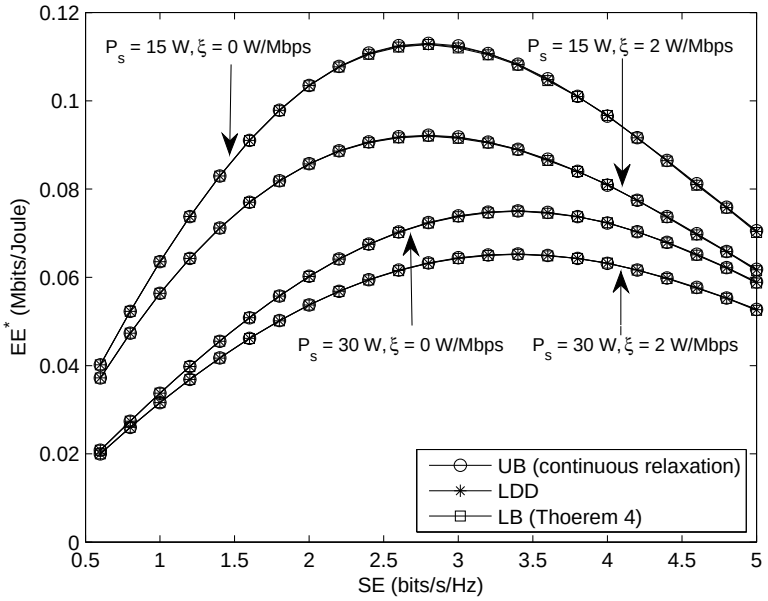
2.4 Numerical Results

In this section, we present some simulation results to verify the theoretical analysis and the effectiveness of the proposed approaches. In our simulation, the frequency spacing between adjacent subcarriers is 15 kHz and the noise power density is -174 dBm/Hz. The frequency-selective Rayleigh fading is generated with the ITU Pedestrian-B model [75].

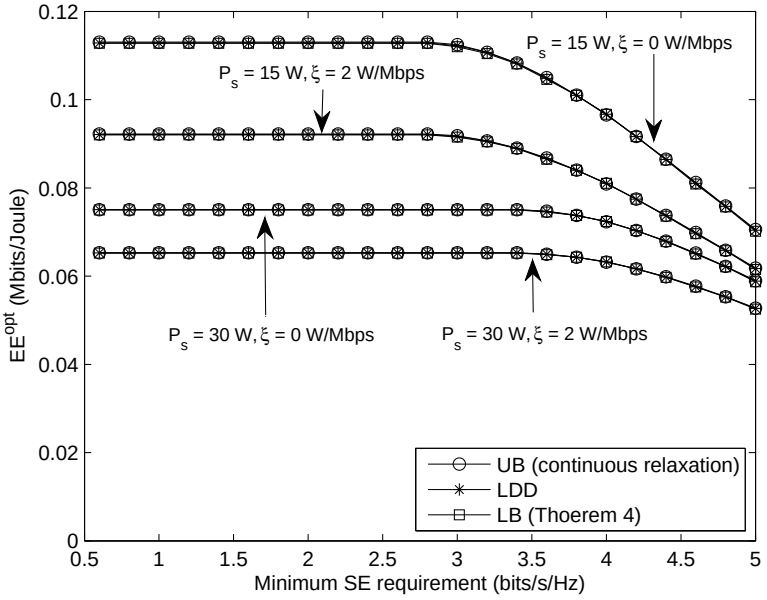
The rate requirements for real-time users are 200 kbps while the minimum rate requirements for non-real-time users are 50 kbps. The fairness notion employed here for non-real-time users is the *partial proportional constraint* (PPC) modified from the *propotional rate constraint* in [72], where $\omega_{K_0} : \omega_{K_0+1} : \cdots : \omega_K = \alpha_{K_0} : \alpha_{K_0+1} : \cdots : \alpha_K$. For simplicity, we let $\alpha_k = \check{R}_k \log_2 \bar{g}_k$, where $\bar{g}_k$ is the geometric mean of the instantaneous channel power gains of user k. Consequently $\omega_k = \check{R}_k \log_2\left(1 + \bar{g}_k\right) / \sum_{k'=K_0}^{K} \check{R}_{k'} \log_2\left(1 + \bar{g}_{k'}\right)$. The drain efficiency of power amplifier is assumed to be 38%.

Figure 2.3 demonstrates the EE-SE relation in the case that two real-time and four non-real-time users with an identical average *channel-gain-to-noise ratio* (CNR) of 20 dB compete for 72 data subcarriers. Different static circuit power (15, 30 W) and dynamic circuit power factors (0, 2 W/Mbps) are considered. No maximum transmit power constraint or specific minimum overall throughput requirement is imposed to investigate the performance limit. From Figure 2.3(a), the EE-SE relation has a bell shape curve and is also quasiconcave, since the upper bound and the lower bound almost perfectly match and they are in a bell shape. Indeed, this quasiconcavity is the foundation of the proposed JIOO methodology. It also infers that the LDD approach can serve as an optimal inner-layer algorithm candidate which leads to the optimal EE-SE relation. Figure 2.3(a) also compares and indicates the influence of circuit power on the EE-SE relation. From there, the η_{EE}^{opt} decreases with the circuit power while the η_{SE}^{opt} increases with the static circuit power but is independent to the dynamic circuit power factor, which is consistent with the Property 2.2. Figure 2.3(b) plots the η_{EE}^{opt}-versus-$\check{\eta}_{SE}$ curves corresponding to curves in Figure 2.3(a). It indicates that for the different minimum throughput requirements, $B\check{\eta}_{SE}$, which are smaller than or equal to a certain threshold, optimal energy-efficient designs lead to a same throughput. However, for the minimum throughput requirement that is larger than the threshold, the most energy-efficient transmission strategy is to operate exactly at the required minimum throughput.

Next, we evaluate the EE, SE, transmit power, and total power of the JIOO algorithms



(a) η_{EE}^{*}-versus-η_{SE} curves with different circuit power parameters

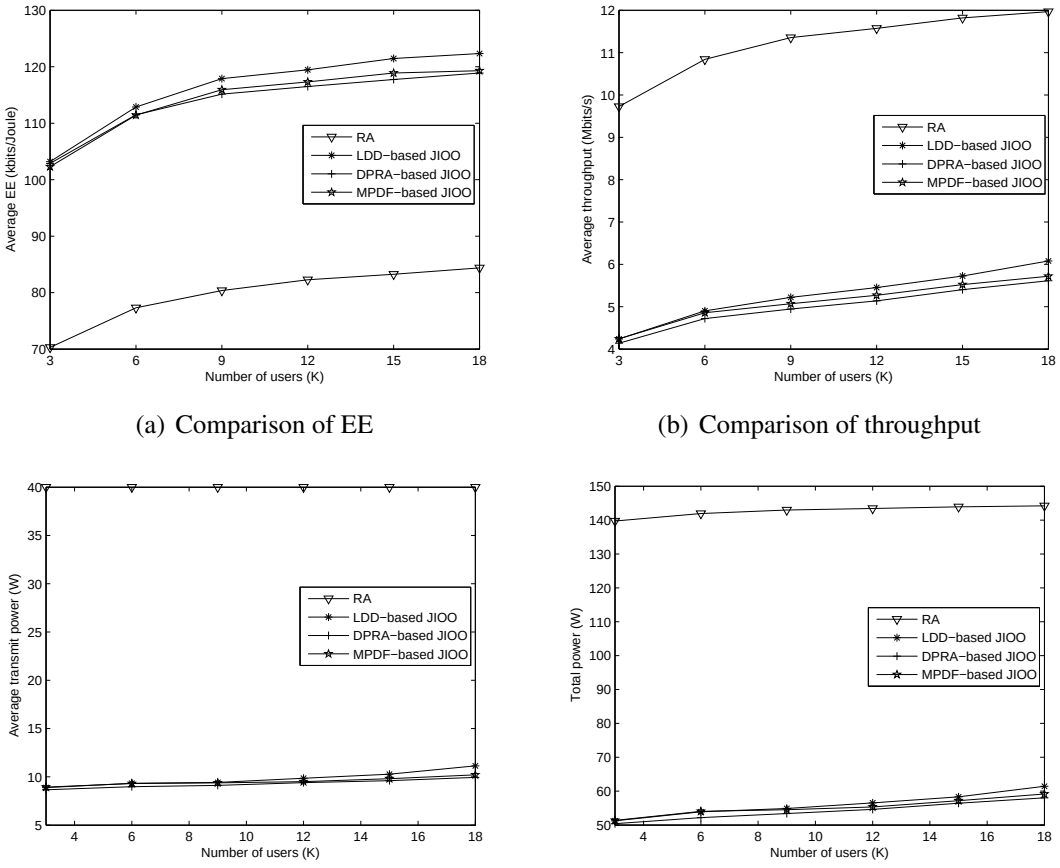


(b) η_{EE}^{opt}-versus-$\check{\eta}_{SE}$ curves with different circuit power parameters

Figure 2.3. EE-SE relation obtained by the continuous relaxation, LDD and LDD-based methods.

with inner-layer algorithm based on the LDD, DPRA and MPDF methods, respectively, and compare them with the conventional RA resource allocation scheme with PPC[2]. In this simulation example, the maximum transmit power is 40 W. The static circuit power is 15 W and the dynamic circuit power factor is 2 W/Mbps. All users still have the same average CNR and the ratio of the number of real-times users to the number of non-real-time users is fixed to be $\frac{1}{2}$ for different number of total users and number of subcarriers. No specific minimum overall throughput requirement is required as long as the per-user QoS is ensured. From Figure 2.4(a) and 4(b), for energy-efficient design, EE and SE both increase with CNR, which is consistent with Property 2.1. From Figures 4(c) and 4(d), energy-efficient design tends to use less transmit power when the CNR improves; and in the lower CNR regime the transmit-related power (ζP) dominates the total power consumption while in the high CNR regime the circuit power ($P_s + \xi R$) does for the energy-efficient design. It can be seen that compared with the spectral-efficient design, the energy-efficient design can achieve higher EE with less transmit power and total power at the cost of some throughput loss. From Figure 2.4, for both cases ($N = 72, K = 6$ and $N = 180, K = 12$), the MPDF-based and DPRA-based JIOO algorithms not only well approach the optimal EE achieved by the LDD-based JIOO algorithm but also have similar throughput, transmit power and total power as the latter. From Figure 2.4, the LDD-based, MPDF-based, and DPRA-based JIOO approaches do not converge to exactly the same operating points with respect to throughput, transmit power and total power. From subfigures 2.4(c) and 2.4(d), the transmit power and total power of the LDD-based JIOO approach is slightly larger than that of the MPDF-based one and that of the DPRA-based one. It is because the LDD-based JIOO approach also has larger throughput than that of the other two, as shown in subfigure 2.4(b).

[2]Here the RA resource allocation scheme with PPC maximizes the throughput by utilizing full transmit power under proportional rate constraint in (15b).



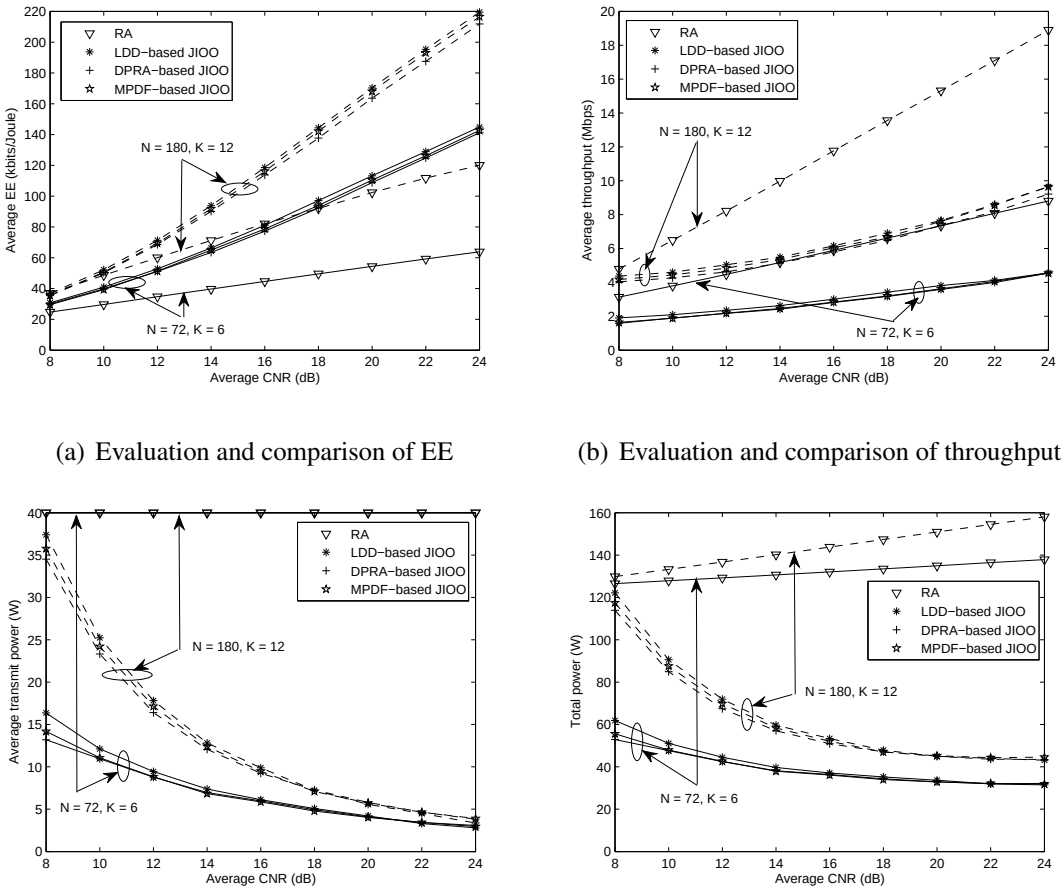
(a) Comparison of EE

(b) Comparison of throughput

(c) Comparison of transmit power

(d) Comparison of total power

Figure 2.4. Performance comparison with different number of users in the case that $N = 180$**,** $|\mathcal{K}_1| : |\mathcal{K}_2| = 1 : 2$**,** $\mathrm{CNR}_k = 16\,\mathrm{dB}$**,** $P_{\max} = 40\,\mathrm{W}$**,** $P_s = 15\,\mathrm{W}$**, and** $\xi = 2\,\mathrm{W/Mbps}$**.**



(a) Evaluation and comparison of EE

(b) Evaluation and comparison of throughput

(c) Evaluation and comparison of transmit power

(d) Evaluation and comparison of total power

Figure 2.5. Performance evaluation and comparison with different CNRs in the case that $N = 72, K = 6$ **and** $N = 180, K = 12$**.**

Figure 2.5 plots the performance with different number of users at a fixed CNR. It can be seen that EE and SE do increase but at a modest rate while transmit power increases very slowly. The explanation of this interesting phenomenon in the example is as follows. Since the probability for the new active users to have better or poorer instantaneous channel quality is equal and (minimum) rate constraints and PPC are required, the benefits from the good users are encumbered by the poor users to a certain extent.

We also include crude average computational time measurements for different methods using the `profile` function in Matlab, which is shown in Figure 2.6. Similar configuration as the one in last example is used. It can be seen that the complexity of the MPDF-based

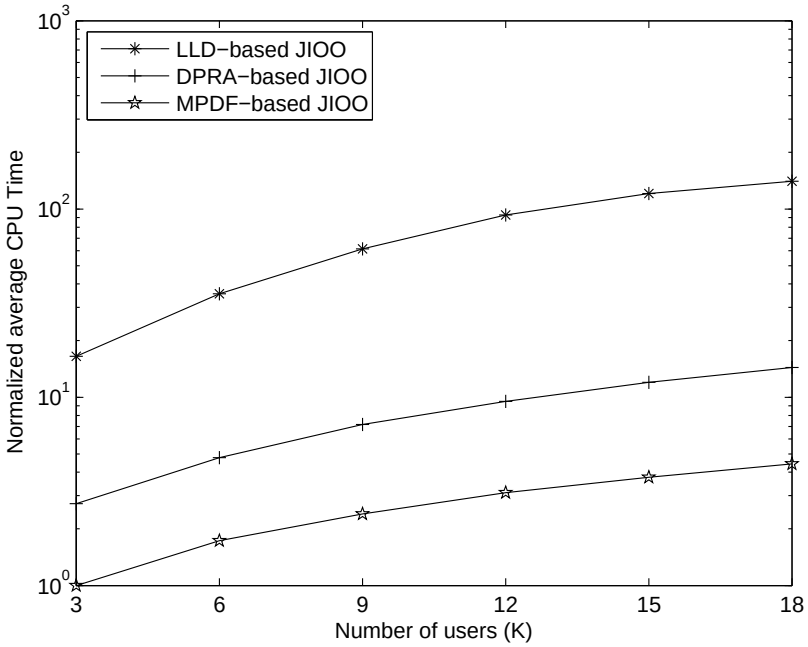


Figure 2.6. Complexity comparison of JIOO algorithms in the case that $N = 180$**,** $|\mathcal{K}_1| : |\mathcal{K}_2| = 1 : 2$**,** $\mathrm{CNR}_k = 16\,\mathrm{dB}$**,** $P_{\max} = 40\,\mathrm{W}$**,** $P_s = 15\,\mathrm{W}$**, and** $\xi = 2\,\mathrm{W/Mbps}$**.**

JIOO algorithm is lower than that of the DPRA-based one and that of the LLD-based one, respectively. In the case with a moderate number of users, i.e., $K = 12$, the complexity of the proposed MPDF-based JIOO algorithm is about 33% and 3% of that of the DPRA-based one and the LLD-based one, respectively. Thus, if the object is EE improvement with minimal complexity, our MPDF-based method offers an attractive performance to complexity good tradeoff.

2.5 Conclusion

In this chapter, we have investigated the EE-SE relation in a single cell downlink OFDMA network, which is important for designing GR networks that require a better balance between EE and SE. The EE-SE relation is proved to be a quasiconcave function, and the impact of channel power gain and circuit power is analyzed. Then a tight lower bound and a tight upper bound on the EE-SE curve are obtained by LDD and continuous relaxation, respectively. For the case with priority and fairness consideration, an alternative easily

derived upper bound and a low-complexity near-optimal resource allocation algorithm are proposed, respectively. Simulation results confirm the theoretical findings and demonstrate that the proposed resource allocation algorithm can efficiently approach the desirable EE-SE tradeoff in practical situation with heterogeneous traffic and consideration of priority and/or fairness.

CHAPTER 3

ENERGY-EFFICIENT RESOURCE ALLOCATION IN OFDMA NETWORKS

In this chapter, we address the energy-efficient resource allocation, including performance limit and optimal and low-complexity suboptimal algorithms, in both downlink and uplink of OFDMA networks with frequency-selective fading channels and consideration of QoS and priority/fainess issues. For the downlink scenario, we model the problem as the maximization of generalized EE under QoS requirements while for the uplink we model the problem as the maximization of the minimum individual EE under QoS requirements. For both cases, we first give the optimal solution then develop a low-complexity suboptimal solution by exploring the inherent structure and property of the energy-efficient design. For the downlink case, we also find a tractable upper bound on EE, which is further demonstrated to be relatively tight when the number of subcarriers is larger than that of UEs and is the foundation of a near-optimal approach.

The rest of this chapter is organized as follows. In Section 3.1, we describe the system model and formulate the optimization problem for both the downlink and uplink cases. In Section 3.2, the downlink scenario is investigated. We give the optimal solution by exhaustive search for reference, find a numerically tractable upper bound on EE and accordingly propose an efficient near-optimal approach, then propose a suboptimal approach by examining the inherent structure of the EE objective function. In Section 3.3, we provide both the optimal and low-complexity suboptimal solutions for the uplink scenario. Then, we present numerical results in Section 3.4. Finally, we conclude this chapter in Section 3.5.

3.1 Problem Description

In this section, we introduce the system model of OFDMA networks and formulate the problems of energy-efficient design.

3.1.1 System Description

We consider a single cell OFDMA network, either downlink or uplink, consisting of K active UEs. The total bandwidth, B, is equally divided into N subcarriers, each with a bandwidth of $W = \frac{B}{N}$.

Assume that each subcarrier is exclusively assigned to at most one UE each time to avoid interference among different UEs. Denote the transmit power and the channel frequency response of the kth UE on the nth subcarrier as $p_{k,n}$ and $H_{k,n}$, respectively. Then, the maximum achievable data rate of the kth UE on the nth subcarrier is accordingly

$$r_{k,n} = W \log_2 \left(1 + \frac{p_{k,n}|H_{k,n}|^2}{N_0 W}\right), \tag{3.1}$$

where N_0 is the single-sided noise spectral density. Then, the aggregate rate for the kth UE and the overall throughput are

$$R_k = \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n} \text{ and } R = \sum_{k\in\mathcal{K}} R_k = \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n},$$

respectively, where $\rho_{k,n} \in \{1, 0\}$ indicates whether or not the nth subcarrier is assigned to the kth UE, $\mathcal{N} = \{1, 2, \cdots, N\}$ and $\mathcal{K} = \{1, 2, \cdots, K\}$ denote the sets of all subcarriers and all UEs, respectively. Obviously, a feasible subcarrier assignment indictor matrix, $\boldsymbol{\rho} = [\rho_{k,n}]_{K\times N}$, should satisfy

$$\boldsymbol{\rho} \in \varrho \triangleq \left\{ [\rho_{k,n}]_{K\times N} | \sum_{k\in\mathcal{K}} \rho_{k,n} \leq 1, \forall\, n \in \mathcal{N}; \rho_{k,n} \in \{0, 1\}, \forall\, k \in \mathcal{K}, n \in \mathcal{N} \right\}. \tag{3.2}$$

The subcarrier assignment constraint in (3.2) can be equivalently viewed from the following perspective

$$\bigcup_{k=1}^{K} \mathcal{S}_k \subseteq \mathcal{N} \text{ and } \mathcal{S}_k \bigcap \mathcal{S}_{k'} = \varnothing, \forall\, k \neq k',$$

where $\mathcal{S}_k$ is the set of subcarriers assigned to the kth UE.

For practical systems, the total transmit power of either base station or UE is nonnegative and also limited. Thus, any possible power allocation matrix, $\boldsymbol{P} = [p_{k,n}]_{K\times N}$, should be

subject to

$$\begin{aligned}\boldsymbol{P} \in \mathcal{P} \triangleq \Big\{ & [p_{k,n}]_{K\times N} | p_{k,n} \geq 0, \forall\, k \in \mathcal{K}, \forall\, n \in \mathcal{N}; \\ & \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}} p_{k,n} \leq P_{\max} \text{ (for downlink case)}; \\ & \sum_{n\in\mathcal{N}} p_{k,n} \leq P_k^{\max}, \forall k \in \mathcal{K} \text{ (for uplink case)} \Big\},\end{aligned} \tag{3.3}$$

where $P_{\max}$ and $P_k^{\max}$ represent the maximum total transmit power at base station for downlink transmission and at the kth UE for uplink transmission, respectively. Then, the overall transmit power for the kth UE and the total transmit power are

$$P_k = \sum_{n\in\mathcal{N}} p_{k,n} \text{ and } P = \sum_{k\in\mathcal{K}} P_k = \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}} p_{k,n},$$

respectively.

Besides transmit power, the energy consumption also includes circuit energy consumption incurred by active circuit blocks, such as analog-to-digital converter, digital-to-analog converter, synthesizer, and mixer [24]. For the downlink transmission, the overall power consumption at the base station is given by [24]

$$P_{tot} = \zeta P + P_c, \tag{3.4}$$

where ζ is the reciprocal of the drain efficiency of power amplifier and P_c represents the circuit power consumption. Similarly, the overall power consumption at the kth uplink UE is modeled as

$$P_k^{tot} = \zeta_k P_k + P_k^c, \tag{3.5}$$

where ζ_k and P_k^c are the corresponding reciprocal of the drain efficiency of power amplifier and circuit power consumption, respectively.

3.1.2 EE for Downlink Transmission

We define the generalized EE for the downlink transmission as the weighted totally delivered bits per unit energy, i.e.,

$$\eta_{EE}^{DL} \triangleq \frac{\sum_{k\in\mathcal{K}} \omega_k R_k}{\zeta P + P_c}, \tag{3.6}$$

where the predetermined weights, ω_k's, can potentially provide certain level of priority and/or fairness among the UEs and the generalized EE still reflects the notion of the conventional EE that is defined as $\frac{\sum_{k\in\mathcal{K}} R_k}{\zeta P + P_c}$ in [3, 26, 37, 28], where $\omega_1 = \omega_2 = \cdots = \omega_K = 1$. To provide different service priorities and guarantee QoS for each UE, we consider the generalized EE under a series of traffic-related minimum rate requirements, $\check{R}_k$'s, and the peak transmit power, $P_{\max}$. The generalized EE optimization problem for the downlink transmission can be mathematically formulated as

$$\hat{\eta}_{EE}^{DL} \triangleq \max_{\boldsymbol{\rho}\in\mathcal{Q},\boldsymbol{P}\in\mathcal{P}} \frac{\sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}}{\sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}} \zeta p_{k,n} + P_c}, \tag{3.7a}$$

subject to

$$\sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n} \geq \check{R}_k, \forall\, k \in \mathcal{K}, \tag{3.7b}$$

where $\hat{\eta}_{EE}^{DL}$ represents the optimal downlink EE.

3.1.3 EE for Uplink Transmission

Arithmetic and geometric average EE have been proposed as optimization objectives for uplink energy-efficient design [27, 38]. In this chapter, we optimize the minimum individual EE instead, which guarantees satisfying EE as much as possible even for the worst UE. The individual EE of the kth UE for the uplink transmission is conventionally defined as

$$\eta_{EE,k}^{UL} \triangleq \frac{R_k}{\zeta_k P_k + P_k^c}, \tag{3.8}$$

like that in [27]. Mathematically, the EE optimization problem for the uplink transmission can be expressed as

$$\hat{\eta}_{EE}^{UL} \triangleq \max_{\boldsymbol{\rho}\in\mathcal{Q},\boldsymbol{P}\in\mathcal{P}} \left\{ \min_{k\in\mathcal{K}} \frac{\sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}}{\sum_{n\in\mathcal{N}} \zeta_k p_{k,n} + P_k^c} \right\} \tag{3.9a}$$

subject to

$$\sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n} \geq \check{R}_k, \forall\, k \in \mathcal{K}. \tag{3.9b}$$

3.2 Downlink Transmission

In this section, we will develop the optimal and low-complexity suboptimal approaches for the energy-efficient resource allocation in the downlink transmission and study the performance limit of such energy-efficient design.

3.2.1 Optimal Solution

Problem (3.7) is in general NP-hard for the optimal solution. To obtain insight on the problem, we first investigate the properties of the case with a given subcarrier assignment, which are summarized in the following theorem and proved in Appendix B.1.

Theorem 3.1 *For any fixed subcarrier assignment indicator matrix $\boldsymbol{\rho} \in \varrho$ and its corresponding subcarrier assignment sets $\mathcal{S}_k$'s ($\forall k \in \mathcal{K}$), the maximum achievable EE at a certain total transmit power, P, namely,*

$$\begin{aligned}\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P) &\triangleq \max_{p_{k,n}\geq 0} \frac{R_w^{(\boldsymbol{\rho})}(\boldsymbol{P})}{\zeta P + P_c} \\ &\triangleq \max_{p_{k,n}\geq 0} \frac{\sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{S}_k} r_{k,n}}{\zeta P + P_c},\end{aligned} \tag{3.10a}$$

subject to

$$\sum_{n\in\mathcal{S}_k} r_{k,n} \geq \check{R}_k, \forall\, k \in \mathcal{K}, \tag{3.10b}$$

$$\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{S}_k} p_{k,n} = P, \tag{3.10c}$$

has the following properties:

(i) $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ is continuously differentiable and strictly quasiconcave in P,

(ii) $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ either strictly decreases or first strictly increases and then strictly decreases with P starting from $P_0 = \sum_{k\in\mathcal{K}} R_k^{-1}\left(\mathcal{S}_k, \check{R}_k\right)$,[1]

[1] $R_k(\mathcal{X}, P_k)$ and $R_k^{-1}(\mathcal{X}, R_k)$ denote the maximum aggregate data rate achieved by optimally allocating total power P_k over subcarrier set $\mathcal{X}$ and the minimum transmit power required for realizing aggregate rate R_k over subcarrier set $\mathcal{X}$ for the kth UE, respectively.

(iii) $\frac{\mathrm{d}\hat{\eta}_{EE}^{(\rho)}(P)}{\mathrm{d}P}$ $\begin{cases} > 0 & \text{if } \hat{\eta}_{EE}^{(\rho)}(P) < \frac{1}{\zeta}\frac{\mathrm{d}\hat{R}_w^{(\rho)}(P)}{\mathrm{d}P} \\ = 0 & \text{if } \hat{\eta}_{EE}^{(\rho)}(P) = \frac{1}{\zeta}\frac{\mathrm{d}\hat{R}_w^{(\rho)}(P)}{\mathrm{d}P} \\ < 0 & \text{if } \hat{\eta}_{EE}^{(\rho)}(P) > \frac{1}{\zeta}\frac{\mathrm{d}\hat{R}_w^{(\rho)}(P)}{\mathrm{d}P} \end{cases}$,

where $\hat{R}_w^{(\rho)}(P) \triangleq \max_{p_{k,n}\geq 0} R_w^{(\rho)}(\boldsymbol{P}) = \max_{p_{k,n}\geq 0} \sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{S}_k} r_{k,n}$ *under constraints (3.10b) and (3.10c) is the maximum Weighted Sum Rate (WSR). And its derivative satisfies*

$$\begin{aligned} \frac{\mathrm{d}\hat{R}_w^{(\rho)}(P)}{\mathrm{d}P} &\equiv \max_{k\in\mathcal{K}, n\in\mathcal{S}_k} \frac{\omega_k W g_{k,n}\log_2 e}{1+\hat{p}_{k,n}g_{k,n}} \\ &\equiv \begin{cases} \max\limits_{k\in\mathcal{K}} \frac{\omega_k W\log_2 e}{\mu_k} & \text{if } P = P_0 \\ \frac{\omega_k W\log_2 e}{\mu} & \text{if } P > P_0 \end{cases}, \end{aligned} \tag{3.11}$$

where $g_{k,n} \triangleq \frac{|H_{k,n}|^2}{N_0 W}$ *is the channel-gain-to-noise ratio (CNR) of the kth UE on the nth subcarrier and* $\hat{p}_{k,n}$ $(n \in \mathcal{S}_k)$ *is the optimal power on the nth subcarrier for achieving* $\hat{R}_w^{(\rho)}(P)$*. The optimal power can be calculated from the following water-filling equations*

$$\check{p}_{k,n} = \left(\mu_k - \frac{1}{g_{k,n}}\right)^+, \forall\, n \in \mathcal{S}_k, \tag{3.12a}$$

$$\sum_{n\in\{n\in\mathcal{S}_k|\check{p}_{k,n}>0\}} W\log_2(\mu_k g_{k,n}) = \check{R}_k. \tag{3.12b}$$

$$\hat{p}_{k,n} = \check{p}_{k,n} + \left(\mu - \frac{1}{g_{k,n}} - \check{p}_{k,n}\right)^+, \tag{3.13a}$$

$$\sum_{k\in\mathcal{K}} \sum_{n\in\{n\in\mathcal{S}_k|\hat{p}_{k,n}>\check{p}_{k,n}\}} \left(\omega_k\mu - \frac{1}{g_{k,n}} - \check{p}_{k,n}\right) = P - \sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{S}_k} \check{p}_{k,n}, \tag{3.13b}$$

where $(x)^+$ *represents* $\max(x, 0)$ *and* μ_k *and* μ *are intermediate variables.*

For any strictly quasiconcave function, there is always a unique global maximum. Thus, Property (i) guarantees the existence and uniqueness of the global maximum and reveals the differentiability of $\hat{\eta}_{EE}^{(\rho)}(P)$. Property (ii) further indicates that the maximum is always achieved at a finite transmit power. Property (iii) connects the sign of the first derivative with the relative size of the EE and the scaled reciprocal of the water-filling level. From

Theorem 3.1, once the subcarrier assignment is fixed, the corresponding optimal power allocation strategy for (8) can be easily obtained by a derivative-assisted bisection method that is based on the single-UE water-filling in (3.12), and the multilevel water-filling in (3.13). The algorithm is named *bisection-based power adaptation* (BPA) and is sketched in Table 3.1. The basic idea of the BPA algorithm is to search bidirectionally for the maximum of $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ by utilizing Properties (i), (ii), and (iii). It starts by doing the single-user water-filling to check the monotonicity of $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ at P_0, which is shown from Line 1 to Line 5. If $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ is decreasing at P_0, the optimal EE is achieved at P_0 and the BPA algorithm terminates; otherwise, it has to find another feasible point at which $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ is decreasing or it finds out that $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ is even increasing at $P_{\max}$, which is described from Line 6 to Line 26. Then, it searches for the optimal total transmit power (the corresponding EE is optimal) between the two boundary points ($P^{(1)}$ and $P^{(2)}$), which is illustrated from Line 27 to Line 37. In a specific case of $K = 1$, i.e., the OFDM case, similar problems as (3.7) have been investigated and solved by the BSAA algorithm in [26], the parameterized convex optimization in [37], and the fixed-point algorithm in [28], respectively. Compared with them, the proposed BPA algorithm simply provides another approach and perspective, which is well connected with the mature water-filling technique and is also easy to implement.

The optimal solution to (3.7) can be obtained by applying the BPA algorithm to every feasible subcarrier assignment $\boldsymbol{\rho} \in \varrho$ and then choose the one with the maximum EE. However, the complexity is extremely high and makes it prohibitive for practical scenarios.

Table 3.1. Bisection-based Power Adaptation (BPA) Algorithm.

Algorithm *BPA*

Input: $\boldsymbol{\rho} = [\rho_{k,n}]_{K\times N}$, P_c, $P_{\max}$; ω_k, $\check{R}_k$, $\forall\ k \in \mathcal{K}$

Output: $\boldsymbol{P} = [p_{k,n}]_{K\times N}$

---

for each UE $k \in \mathcal{K}$

  Do single-user water-filling using (3.12) to get $\check{p}_{k,n}$ and μ_k;

end

$\boldsymbol{P}^{(1)} \triangleq [p^{(1)}_{k,n}]_{K\times N} \leftarrow [\check{p}_{k,n}]_{K\times N}$; $P^{(1)} \leftarrow \sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{S}_k} p^{(1)}_{k,n}$;

$\hat{\eta}^{(\boldsymbol{\rho}),(1)}_{EE} \leftarrow \frac{\sum_{k\in\mathcal{K}}\omega_k\check{R}_k}{\zeta P^{(1)}+P_c}$; $d^{(1)} \leftarrow \max\limits_{k\in\mathcal{K}} \frac{\omega_k W \log_2 e}{\mu_k}$;

if $\hat{\eta}^{(\boldsymbol{\rho}),(1)}_{EE} \geq \frac{1}{\zeta} d^{(1)}$

  $\boldsymbol{P} \triangleq [p_{k,n}]_{K\times N} \leftarrow \boldsymbol{P}^{(1)}$;

  return;

else

  $\boldsymbol{P}^{(2)} \triangleq [p^{(2)}_{k,n}]_{K\times N} \leftarrow \boldsymbol{P}^{(1)}$; $P^{(2)} \leftarrow P^{(1)}$;

  $P^{(1)} \leftarrow \min\left(\kappa P^{(1)}, P_{\max}\right)$, where $\kappa > 1$, e.g., $\kappa \leftarrow 1.5$;

  Do multilevel water-filling with total transmit power $P^{(1)}$

    using (3.13) to get $\boldsymbol{P}^{(1)} \leftarrow [\hat{p}_{k,n}]_{K\times N}$ and μ;

  $\hat{\eta}^{(\boldsymbol{\rho}),(1)}_{EE} \leftarrow \frac{\sum_{k\in\mathcal{K}}\omega_k\sum_{n\in\mathcal{S}_k}\log_2\left(1+p^{(1)}_{k,n}g_{k,n}\right)}{\zeta P^{(1)}+P_c}$; $d^{(1)} \leftarrow \frac{\omega_k W log_2 e}{\mu}$;

  while $\hat{\eta}^{(\boldsymbol{\rho}),(1)}_{EE} < \frac{1}{\zeta} d^{(1)}$ && $P^{(1)} < P_{\max}$

    $\boldsymbol{P}^{(2)} \leftarrow \boldsymbol{P}^{(1)}$; $P^{(2)} \leftarrow P^{(1)}$;

    $P^{(1)} \leftarrow \min\left(\kappa P^{(1)}, P_{\max}\right)$;

    Do multilevel water-filling with $P^{(1)}$ using (3.13) to get

      $\boldsymbol{P}^{(1)} \leftarrow [\hat{p}_{k,n}]_{K\times N}$ and μ;

    $\hat{\eta}^{(\boldsymbol{\rho}),(1)}_{EE} \leftarrow \frac{\sum_{k\in\mathcal{K}}\omega_k\sum_{n\in\mathcal{S}_k}\log_2\left(1+p^{(1)}_{k,n}g_{k,n}\right)}{\zeta P^{(1)}+P_c}$; $d^{(1)} \leftarrow \frac{\omega_k W log_2 e}{\mu}$;

  end

  if $\hat{\eta}^{(\boldsymbol{\rho}),(1)}_{EE} \leq \frac{1}{\zeta} d^{(1)}$

    $\boldsymbol{P} \triangleq [p_{k,n}]_{K\times N} \leftarrow \boldsymbol{P}^{(1)}$;

    return;

  end

end

while no convergence

  $P \leftarrow \frac{P^{(1)}+P^{(2)}}{2}$;

  Do multilevel water-filling with P using (3.13) to get

    $\boldsymbol{P} \triangleq [p_{k,n}]_{K\times N} \leftarrow [\hat{p}_{k,n}]_{K\times N}$ and μ;

  $\hat{\eta}^{(\boldsymbol{\rho})}_{EE} \leftarrow \frac{\sum_{k\in\mathcal{K}}\omega_k\sum_{n\in\mathcal{S}_k}\log_2\left(1+p_{k,n}g_{k,n}\right)}{\zeta P+P_c}$; $d \leftarrow \frac{\omega_k W log_2 e}{\mu}$;

  if $\hat{\eta}^{(\boldsymbol{\rho})}_{EE} < \frac{1}{\zeta} d$

    $\boldsymbol{P}^{(2)} \leftarrow \boldsymbol{P}$; $P^{(2)} \leftarrow P$;

  else

    $\boldsymbol{P}^{(1)} \leftarrow \boldsymbol{P}$; $P^{(1)} \leftarrow P$;

  end

end

3.2.2 Near-Optimal Solution

To facilitate practical application of the optimal energy-efficient design, we will first exploit and prove the quasiconcave relation between an upper bound on $\max_{\boldsymbol{\rho}\in\mathcal{Q}} \hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ and total transmit power, P.

Define $\tilde{\boldsymbol{\rho}}$, $\tilde{\mathcal{Q}}$, $\tilde{\rho}_{k,n}$, and $\tilde{r}_{k,n}$ as follows

$$\begin{aligned}\tilde{\boldsymbol{\rho}} \in \tilde{\mathcal{Q}} \triangleq \Big\{[\tilde{\rho}_{k,n}]_{K\times N} | \sum_{k\in\mathcal{K}} \tilde{\rho}_{k,n} \leq 1, \forall\, n \in \mathcal{N};\\ \tilde{\rho}_{k,n} \in [0,1], \forall\, k \in \mathcal{K}, n \in \mathcal{N}\Big\}.\end{aligned} \tag{3.14}$$

$$\tilde{r}_{k,n} = W \log_2\left(1 + \frac{p_{k,n}|H_{k,n}|^2}{\tilde{\rho}_{k,n} N_0 W}\right). \tag{3.15}$$

Note that when $\tilde{\rho}_{k,n}$ approaches zero, $\tilde{\rho}_{k,n}\tilde{r}_{k,n}$ also tends to be zero (although $\tilde{r}_{k,n}$ goes to infinity), which agrees with that the nth subcarriers is nearly not assigned to the kth UE. When $\tilde{\rho}_{k,n}$ is close to one, $\tilde{\rho}_{k,n}\tilde{r}_{k,n}$ is close to $\rho_{k,n}r_{k,n}$, which indicates that the nth subcarrier is almost entirely assigned to the kth UE. Therefore, when $\tilde{\rho}_{k,n}$ is close to zero or one, the approximation of $\rho_{k,n}r_{k,n}$ by $\tilde{\rho}_{k,n}\tilde{r}_{k,n}$ becomes precise. Moreover, $\tilde{\rho}_{k,n}\tilde{r}_{k,n}$ is easier to deal with mathematically compared to $\rho_{k,n}r_{k,n}$. As a result of its nice tractability and acceptable accuracy, such approximation is widely used in literature on OFDMA resource allocation [70, 71, 69, 46]. We will use $\tilde{\rho}_{k,n}\tilde{r}_{k,n}$ instead of $\rho_{k,n}r_{k,n}$ to represent the rate of the kth UE on the nth subcarrier in the following modified EE optimization problem.

We then formulate a modified EE optimization problem of (3.7) as follows

$$\hat{\eta}_{EE}^{UB} \triangleq \max_{\tilde{\boldsymbol{\rho}}\in\tilde{\mathcal{Q}}, \boldsymbol{P}\in\mathcal{P}} \frac{\sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n}\tilde{r}_{k,n}}{\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}} \zeta p_{k,n} + P_c}, \tag{3.16a}$$

subject to

$$\sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n}\tilde{r}_{k,n} \geq \breve{R}_k, \forall\, k \in \mathcal{K}. \tag{3.16b}$$

In (3.16), such fractional $\tilde{\rho}_{k,n}$'s can be either interpreted as frequency domain sharing of subcarriers [70, 71] or regarded as time domain sharing of subcarriers [69]. Clearly, by

relaxing $\rho_{k,n} \in \{0,1\}$ to $\tilde{\rho}_{k,n} \in [0,1]$ and replacing $r_{k,n}$ with $\tilde{r}_{k,n}$, problem (3.16) always yields an upper bound on the EE of (3.7), i.e., $\hat{\eta}_{EE}^{UB} \geq \hat{\eta}_{EE}^{DL}$, although it does not necessarily guarantee a solution where $\tilde{\rho}_{k,n}$ is either 0 or 1 and thus the satisfaction of (3.16b) cannot ensure the feasibility of (3.7b). The properties of the EE upper bound are summarized in Theorem 2 and proved in Appendix B.2.

Theorem 3.2 *The upper bound on the maximum achievable EE,* $\max_{\boldsymbol{\rho}\in\mathcal{Q}} \hat{\eta}_{EE}^{(\rho)}(P)$*, at a certain total transmit power, P, namely,*

$$\begin{aligned}\hat{\eta}_{EE}^{UB}(P) &\triangleq \max_{\tilde{\boldsymbol{\rho}}\in\tilde{\mathcal{Q}}, p_{k,n}\geq 0} \frac{\mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P})}{\zeta P + P_c} \\ &\triangleq \max_{\tilde{\boldsymbol{\rho}}\in\tilde{\mathcal{Q}}, p_{k,n}\geq 0} \frac{\sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n}\tilde{r}_{k,n}}{\zeta P + P_c},\end{aligned} \tag{3.17a}$$

subject to

$$\sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n}\tilde{r}_{k,n} \geq \check{R}_k, \forall\, k \in \mathcal{K}, \tag{3.17b}$$

$$\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}} p_{k,n} = P, \tag{3.17c}$$

has the following properties

(i) $\hat{\eta}_{EE}^{UB}(P)$ *is continuously differentiable and strictly quasiconcave in P,*

(ii) $\hat{\eta}_{EE}^{UB}(P)$ *either strictly decreases or first strictly increases and then strictly decreases with P,*

(iii) $\frac{\mathrm{d}\hat{\eta}_{EE}^{UB}(P)}{\mathrm{d}P}$
$$\begin{cases} > 0 & \text{if } \hat{\eta}_{EE}^{UB}(P) < \frac{1}{\zeta}\frac{\mathrm{d}\hat{\mathcal{R}}_w(P)}{\mathrm{d}P}, \\ = 0 & \text{if } \hat{\eta}_{EE}^{UB}(P) = \frac{1}{\zeta}\frac{\mathrm{d}\hat{\mathcal{R}}_w(P)}{\mathrm{d}P}, \\ < 0 & \text{if } \hat{\eta}_{EE}^{UB}(P) > \frac{1}{\zeta}\frac{\mathrm{d}\hat{\mathcal{R}}_w(P)}{\mathrm{d}P}, \end{cases}$$

where $\hat{\mathcal{R}}_w(P) \triangleq \max_{\tilde{\boldsymbol{\rho}}\in\tilde{\mathcal{Q}}, p_{k,n}\geq 0} \mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P}) = \max_{\tilde{\boldsymbol{\rho}}\in\tilde{\mathcal{Q}}, p_{k,n}\geq 0} \sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n}\tilde{r}_{k,n}$ *under constraints (3.17b) and (3.17c) is the maximum Modified Weighted Sum Rate (MWSR). And its derivative satisfies*

$$\frac{\mathrm{d}\hat{\mathcal{R}}_w(P)}{\mathrm{d}P} \equiv \max_{k\in\mathcal{K}, n\in\mathcal{N}} \frac{\omega_k W \tilde{\rho}_{k,n}^* g_{k,n} \log_2 e}{1 + \tilde{\rho}_{k,n}^* g_{k,n} p_{k,n}^*}, \tag{3.18}$$

where $\tilde{\boldsymbol{\rho}}^* = [\tilde{\rho}^*_{k,n}]_{K\times N}$ *and* $\boldsymbol{P}^* = [p^*_{k,n}]_{K\times N}$ *are the optimal subcarrier and power allocation matrices for achieving* $\hat{\mathcal{R}}_w(P)$*.*

Theorem 3.2 demonstrates the quasiconcavity of $\hat{\eta}^{UB}_{EE}(P)$, which is an upper bound on the EE that is defined as $\hat{\eta}^{DL}_{EE}(P) \triangleq \max_{\boldsymbol{\rho}\in\mathcal{Q}} \hat{\eta}^{(\boldsymbol{\rho})}_{EE}(P)$ under constraints (3.7b) and (3.17c), in the transmit power, P, and implies the existence and the uniqueness of the global maximum, i.e., $\hat{\eta}^{UB}_{EE}$. More importantly, as a result of the quasiconcavity, problem (3.16) can be decomposed into two layers and solved iteratively by the joint inner- and outer-layer optimization as follows

(i) Inner layer: For a given transmit power, $P \le P_{\max}$, find the maximum EE, $\hat{\eta}^{UB}_{EE}(P)$, and (the sign of) its derivative, $\frac{\mathrm{d}\hat{\eta}^{UB}_{EE}(P)}{\mathrm{d}P}$.

(ii) Outer layer: Search for the transmit power that results in the maximum, $\hat{\eta}^{UB}_{EE}$, by bisection power search like the bisection power search in the BPA algorithm.

The bisection search in the outer-layer is clear and easy. The key lies in the inner-layer algorithm that finds $\hat{\eta}^{UB}_{EE}(P)$ and (the sign of) $\frac{\mathrm{d}\hat{\eta}^{UB}_{EE}(P)}{\mathrm{d}P}$. For a given total transmit power, the inner-layer subproblem to find $\hat{\eta}^{UB}_{EE}(P)$ is equivalent to maximizing the constrained MWSR, which means solving for $\hat{\mathcal{R}}_w(P)$. Since $\mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P}) \triangleq \sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n}\tilde{r}_{k,n}$ is proved to be strictly and jointly concave in $\tilde{\rho}_{k,n}$ and $p_{k,n}$ [70], and the constraint set is convex from Appendix B, the constrained MWSR maximization is in the standard form of a convex programming problem that can be solved by standard numerical methods such as the interior-point method [76]. When the optimal subcarrier assignment matrix, $\tilde{\boldsymbol{\rho}}^*$, and power allocation matrix, $\boldsymbol{P}^*$, for the constrained MWSR maximization problem are obtained, the sign of the first derivative, $\frac{\mathrm{d}\hat{\eta}^{UB}_{EE}(P)}{\mathrm{d}P}$, can be readily determined following Property (iii) in Theorem 3.2. Hence, problem (3.16) can be successfully solved by the aforementioned joint inner- and outer-layer optimization.

When $\hat{\eta}^{UB}_{EE}$ is found, the corresponding optimal $\tilde{\rho}^{opt}_{k,n}$'s are not ensured to be either 0 or 1. To get a feasible solution to the original downlink EE maximization problem (3.7), we need

to round the possibly fractional $\tilde{\rho}_{k,n}^{opt}$'s to 0 or 1 and then perform the BPA algorithm to get the maximum EE for the round-off $\tilde{\rho}_{k,n}^{opt}$'s. Such manipulations may not result in the optimal solution to (3.7). Luckily, this is rarely a problem when the number of subcarriers is large compared with the number of UEs and in this case $\hat{\eta}_{EE}^{UB}$ is quite close to $\hat{\eta}_{EE}^{DL}$. In fact, the optimal $\tilde{\rho}_{k,n}$'s for the constrained MWSR maximization problem mostly tend to be either 0 or 1 when $K \ll N$ [70, 71]. On the other hand, such fine tightness of the EE upper bound and the fact that the optimal $\tilde{\rho}_{k,n}$'s are almost either 0 or 1 implicitly enable the use of the original WSR, $\sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}$, instead of the constrained MWSR for maximization in the inner-layer optimization with an expectation of good performance. This enables us to precisely solve the original problem (3.7) by the joint inner- and outer-layer optimization framework with the constrained WSR maximization as the inner-layer subproblem. The discovery of such fact connects the research on the energy-efficient design to the previous research on the WSR maximization, such as the Lagrange dual decomposition [77] and the branch-and-bound method [78], by allowing them to be inner-layer candidate approaches.

3.2.3 Low-Complexity Suboptimal Solution

Although the convex programming is numerically stable, its computational complexity depends on the number of optimizing variables, which can be large if the number of subcarriers and/or the number of UEs are/is large. Each inner-layer MWSR maximization needs at least $O(NK(1/\delta^2))$ times of water-filling for δ-optimality [67], i.e., $\mathcal{R}_w(P) - \hat{\mathcal{R}}_w(P) < \delta$, where $\mathcal{R}_w(P)$ is the corresponding solution of the convex programming. And the total complexity of (3.16) also depends on the number of iterations in the outer-layer, N_{OL}, and is $O(N_{OL}NK(1/\delta^2))$. Here, we will first explore the inherent property of (3.7) and then, based on this property, a novel low-complexity suboptimal algorithm is proposed to solve problems like (3.7).

Theorem 3.3 *The optimal EE, $\hat{\eta}_{EE}^{DL}$, in (3.7) is always equal to*

$$\hat{\eta}_{EE}^{DL} \equiv \max_{\boldsymbol{\rho}\in\varrho, \boldsymbol{P}\in\mathcal{P}, \boldsymbol{\alpha}\in\alpha} \left\{ \min_{k\in\mathcal{K}} \frac{\omega_k \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}}{\sum_{n\in\mathcal{N}} \zeta p_{k,n} + \alpha_k P_c} \right\}, \tag{3.19a}$$

subject to

$$\sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n} \geq \check{R}_k, \forall\, k \in \mathcal{K}, \tag{3.19b}$$

where $\alpha \triangleq \left\{[\alpha_k]_{K\times 1} | \sum_{k\in\mathcal{K}} \alpha_k = 1; \alpha_k \in \mathbb{R}\right\}$.

Theorem 3.3, which is proved in Appendix B.3, illustrates the structure of the optimal solution in a split form. From it, Corollary 3.1 is readily obtained.

Corollary 3.1 *For any fixed* $\boldsymbol{\alpha} \in \alpha$*, the optimal EE,* $\hat{\eta}_{EE}^{DL}$*, in (3.7), is lower bounded by*

$$\hat{\eta}_{EE}^{DL} \geq \max_{\boldsymbol{\rho}\in\varrho, \boldsymbol{P}\in\mathcal{P}} \left\{ \min_{k\in\mathcal{K}} \frac{\omega_k \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}}{\sum_{n\in\mathcal{N}} \zeta p_{k,n} + \alpha_k P_c} \right\}, \tag{3.20a}$$

subject to

$$\sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n} \geq \check{R}_k, \forall\, k \in \mathcal{K}. \tag{3.20b}$$

Assume $\boldsymbol{\alpha}^{opt} = [\alpha_k^{opt}]_{K\times 1}$ corresponds to the optimal EE, $\hat{\eta}_{EE}^{DL}$, in (3.7) and (3.19). Then, α_k^{opt} can be intuitively regarded as the portion of static circuit power incurred individually by the kth UE when the maximum EE is achieved.[2] Based on (3.19) and (3.20), instead of directly optimizing the EE, we can alternatively maximize the minimum individual EE, i.e., to maximize the objective function $\min_{k\in\mathcal{K}} \frac{\omega_k \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}}{\sum_{n\in\mathcal{N}} \zeta p_{k,n} + \alpha_k P_c}$, under a certain properly chosen $\boldsymbol{\alpha} \in \alpha$, and expect a satisfying EE.

This idea enables us to split one joint and complex optimization objective in (3.7) into a series of relatively isolated and simple objectives in (3.19). It makes subcarrier assignment easier using heuristic algorithms because, for the kth UE, the maximum of its individual EE, $\eta_{EE,k}^{DL} \triangleq \frac{\omega_k \sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}}{\sum_{n\in\mathcal{N}} \zeta p_{k,n} + \alpha_k P_c}$, depends only on its own parameters and the subcarriers it will occupy but not on the power adaptation strategies of other UEs. Here, we propose a greedy subcarrier assignment approach, named, *maximizing-EE-lower-bound-based downlink subcarrier assignment* (MDSA) algorithm, which is motivated by the algorithms in

[2] α_k's may be negative. This is because $\alpha_k P_c$ is only the "virtual" static circuit power caused by the kth UE, which is our intuition/imagination but not necessarily the physical fact. The actual total static circuit power is always P_c.

[71, 72] and is sketched in Table 3.2. The key idea of the MDSA algorithm is to iteratively assign the subcarriers aiming at maximizing the minimum individual EE, $\min_{k\in\mathcal{K}} \eta_{EE,k}^{DL}$, under the QoS requirement. At the beginning, each UE is only virtually assigned its worst subcarrier and the individual EE in this situation is individually optimized under the QoS requirement by the single-UE BPA algorithm and will be used as a benchmark to measure how urgent a UE needs a real subcarrier, which is depicted from Line 1 to Line 4. Then, in each iteration, the UE with the minimum individual EE takes its most favorite subcarrier among all unassigned ones into its exclusively occupied subcarrier set and maximizes its individual EE under the QoS requirement by the single-UE BPA algorithm. The above iteration process will proceed until all subcarriers have been assigned, which is described from Line 5 to Line 10. Besides, no total transmit power constraint is imposed in the MDSA algorithm but it will quite likely be spontaneously satisfied when the subcarrier assignment is finished. This is because: first, with properly chosen $\boldsymbol{\alpha}$, the eventually optimized individual EEs will tend to be of close values.[3] If at a certain stage, one UE requires too much power to guarantee its QoS, then it is likely with a relatively low individual EE and will ask for more subcarriers later to lower its transmit power and increase its individual EE. Second, when the EE gain of the energy-efficient design over the spectral-efficient design is satisfactory, the actually used transmit power for the energy-efficient design should be much less than the maximum available transmit power.

With the promising subcarrier assignment obtained by the MDSA algorithm, we can make further improvement by using the BPA algorithm to find the corresponding EE-optimal power adaptation strategy. The complexity of the MDSA algorithm for a given $\boldsymbol{\alpha}$ is roughly $O(N_{OL}N)$ times of water-filling.

We then suggest an effective way to determine the initial $\boldsymbol{\alpha}$ for the MDSA algorithm. Let $\bar{g}_k \triangleq \mathbb{E}_n(g_{k,n})$ be the average CNR of the kth UE and N_k be the number of subcarriers assigned to the kth UE. We deliberately regard that each UE undergoes flat fading with a

[3] From the proof of Theorem 3, $\boldsymbol{\alpha}^{opt}$ makes all the individual EEs equal when $\hat{\eta}_{EE}^{DL}$ is achieved.

Table 3.2. Maximizing-EE-Lower-Bound-Based Downlink Subcarrier Assignment (MDSA) Algorithm.

Algorithm *MDSA*

Input: $\boldsymbol{\rho} = [\rho_{k,n}]_{K\times N} \leftarrow \mathbf{0}_{K\times N}$; $\mathcal{S}_k \leftarrow \varnothing$, $\forall\, k \in \mathcal{K}$; $\boldsymbol{\alpha} \leftarrow \boldsymbol{\alpha}^{ini}$

Output: $\boldsymbol{\rho}$

1. **for** each UE $k \in \mathcal{K}$
2. Find the subcarrier $\check{n}_k \leftarrow \arg\min_{n\in\mathcal{N}} g_{k,n}$ and calculate
3. $\eta_{EE,k}^{DL} \leftarrow \max\limits_{P_k \geq R_k^{-1}(\{\check{n}_k\},\check{R}_k)} \frac{R_k(\{\check{n}_k\},P_k)}{\zeta P_k + \alpha_k P_c}$;
4. **end**
5. **while** $\mathcal{N} \neq \varnothing$
6. Find the UE $\check{k} \in \mathcal{K}$ such that $\check{k} \leftarrow \arg\min\limits_{k\in\mathcal{K}} \eta_{EE,k}^{DL}$;
7. Find the subcarrier $\hat{n}_{\check{k}} \in \mathcal{N}$ such that $\hat{n}_{\check{k}} \leftarrow \arg\max\limits_{n\in\mathcal{N}} g_{\check{k},n}$;
8. Set $\rho_{\check{k},\hat{n}_{\check{k}}} \leftarrow 1$; $\mathcal{S}_{\check{k}} \leftarrow \mathcal{S}_{\check{k}} \cup \{\hat{n}_{\check{k}}\}$; $\mathcal{N} \leftarrow \mathcal{N} \setminus \{\hat{n}_{\check{k}}\}$;
9. Calculate $\eta_{EE,\check{k}}^{DL} \leftarrow \max\limits_{P_{\check{k}} \geq R_{\check{k}}^{-1}(\mathcal{S}_{\check{k}},\check{R}_{\check{k}})} \left\{\frac{R_k(\mathcal{S}_{\check{k}},P_{\check{k}})}{\zeta P_{\check{k}} + \alpha_{\check{k}} P_c}\right\}$;
10. **end**

CNR of $\bar{g}_k$ and solve the following energy-efficient resource allocation problem

$$\bar{\eta}_{EE}^{DL} \triangleq \max_{N_k\in\mathbb{N},P_k>0} \frac{\sum_{k\in\mathcal{K}} \omega_k N_k W \log_2\left(1 + \frac{\bar{g}_k P_k}{N_k}\right)}{\sum_{k\in\mathcal{K}} \zeta P_k + P_c}, \tag{3.21a}$$

subject to

$$N_k W \log_2\left(1 + \frac{\bar{g}_k P_k}{N_k}\right) \geq \check{R}_k, \forall\, k \in \mathcal{K}, \tag{3.21b}$$

$$\sum_{k\in\mathcal{K}} N_k = N, \tag{3.21c}$$

$$\sum_{k\in\mathcal{K}} P_k \leq P_{\max}. \tag{3.21d}$$

By relaxing N_k's from positive integers to positive real numbers, similar quasiconcave relation can be proved and thus problem (3.21) can be precisely solved by the aforementioned joint inner- and outer-layer optimization framework like (3.16). For such a strictly concave inner-layer MWSR maximization problem, standard convex programming technique such as the interior-point method can be applied for solution [76]. Let $\bar{N}_k$ and $\bar{P}_k$ be the optimal solution to (3.21), then according to (3.19) the initial $\boldsymbol{\alpha}^{ini} = [\alpha_k^{ini}]_{K\times 1}$ is given

by

$$\alpha_k^{ini} = \frac{\omega_k \bar{N}_k W \log_2\left(1 + \frac{\bar{g}_k \bar{P}_k}{\bar{N}_k}\right)}{\bar{\eta}_{EE}^{DL} P_c} - \frac{\zeta \bar{P}_k}{P_c}, \forall\, k \in \mathcal{K}. \tag{3.22}$$

The advantage of this kind of estimation is that (3.21) is much easier to solve than (3.7) and (3.16) because the number of variables is only linear in (two times of) the number of UEs but does not depend on the number of subcarriers, and $\boldsymbol{\alpha}^{ini}$ does not need to update (in the long run) until any of $\bar{g}_k$'s changes (much). The complexity of obtaining $\boldsymbol{\alpha}^{ini}$ is not more than $O(N_{OL}K(1/\delta^2))$ [67, 76].

In Table 3.3, the complexity of the aforementioned optimal, near-optimal, and low-complexity alternative is listed for comparison.

Table 3.3. Complexity Comparison for Downlink Transmission.

Algorithm	Complexity
optimal: brute-force search based on BPA	$O(N_{OL}K^N)$
near-optimal: JIOO based on convex programming	$O(\frac{1}{\delta^2}N_{OL}NK)$
suboptimal: MDSA + BPA	$O(\frac{1}{\delta^2}N_{OL}K + N_{OL}N + N_{OL})$

3.3 Uplink Transmission

In this section, we will discuss the uplink EE optimization.

3.3.1 Optimal Solution

Like the downlink case, the uplink problem (3.9) is, in general, a complicated integer programming problem. However, once the subcarrier assignment is fixed, the maximization of EE for each UE can be regarded as a single-UE case of (3.7) and be readily solved by the BPA algorithm. By exhaustively searching all feasible subcarrier assignments, the optimal solution to (3.9) can be found by choosing the one with the maximum $\min\limits_{k\in\mathcal{K}} \frac{\sum_{n\in\mathcal{N}} \rho_{k,n} r_{k,n}}{\sum_{n\in\mathcal{N}} \zeta_k \rho_{k,n} p_{k,n} + P_{c,k}}$. The total complexity is about $O(KN_{OL}K^N) = O(N_{OL}K^{N+1})$ times of water-filling. Nevertheless, this kind of brute-force method is usually too computationally expensive to afford.

3.3.2 Low-Complexity Suboptimal Solution

The formulation of (3.9) appears very similar to that of (3.19) except that the circuit power is now naturally consumed by each UE individually (α is not introduced) and the transmit power limit for each UE is now separately imposed. Therefore, joint power allocation across different users are not needed any more. Hence, problem (3.9) can be similarly solved by the MDSA with some modification. The modified subcarrier assignment approach, named, *maximizing-minimum-EE-based uplink subcarrier assignment* (MUSA) algorithm, is illustrated in Table 3.4. The basic idea of the MUSA algorithm is to iteratively assign the subcarriers aiming at maximizing the minimum individual EE, $\min_{k\in\mathcal{K}} \eta_{EE,k}^{UL}$, under both the QoS requirement and the transmit power constraint. At first, each UE is also virtually assigned its worst subcarrier for an estimate of its initial EE. For UEs that are not yet capable of supporting their QoS requirements, their initial EEs are obtained by transmitting at their respective maximum transmit power. For UEs that are able to satisfy their QoS requirements, they maximize their EEs under their respective QoS requirements and transmit power constraints by the single-UE BPA algorithm and use them as their initial EEs. Such an initialization process is depicted from Line 1 to Line 8. Then, in each iteration (till all subcarriers are assigned), only the UE with the minimum EE is assigned its most favorite subcarrier among all unassigned ones for an expected EE improvement. To ensure that each UE can eventually meet its QoS requirement, the subcarriers are firstly assigned among the set of UEs that are not yet able to meet their QoS requirements with their exclusively occupied subcarriers and affordable transmit power. When all UEs are capable of guaranteeing their QoS with their occupied subcarriers and affordable transmit power, the remaining subcarriers (if any exists) are then iteratively assigned among the set of all UEs. This part is depicted from Line 9 to Line 26. In Table 3.5, the complexity of the aforementioned optimal and low-complexity suboptimal solutions is listed for comparison.

Table 3.4. Maximizing-minimum-EE-based Uplink Subcarrier Assignment (MUSA) Algorithm.

Algorithm *MUSA*

Input: $\boldsymbol{\rho} = [\rho_{k,n}]_{K\times N} \leftarrow \mathbf{0}_{K\times N}$; $\mathcal{S}_k \leftarrow \varnothing$, $\forall\, k \in \mathcal{K}$; $\mathcal{F} \leftarrow \mathcal{K}$, where $\mathcal{F}$ is the set of UEs that are not yet capable of meeting their QoS requirements with their occupied subcarriers and affordable transmit power

Output: $\boldsymbol{\rho}$

---

for each UE $k \in \mathcal{K}$
  Find the subcarrier $\check{n}_k \in \mathcal{N}$ such that $\check{n}_k \leftarrow \arg\min_{n\in\mathcal{N}} g_{k,n}$;
  if $R_k\left(\{\check{n}_k\}, P_k^{\max}\right) < \check{R}_k$
    $\eta_{EE,k}^{UL} \leftarrow \frac{R_k(\{\check{n}_k\},P_k^{\max})}{\zeta_k P_k^{\max}+P_k^c}$;
  else
    $\eta_{EE,k}^{UL} \leftarrow \max_{R_k^{-1}(\{\check{n}_k\},\check{R}_k)\le P_k\le P_k^{max}} \frac{R_k(\{\check{n}_k\},P_k)}{\zeta_k P_k+P_k^c}$;
  end
end
while $\mathcal{N} \neq \varnothing$
  if $\mathcal{F} \neq \varnothing$
    Find the UE $\check{k} \in \mathcal{F}$ such that $\check{k} \leftarrow \arg\min_{k\in\mathcal{F}} \eta_{EE,k}^{UL}$;
    Find the subcarrier $\hat{n}_{\check{k}} \in \mathcal{N}$ such that $\hat{n}_{\check{k}} \leftarrow \arg\max_{n\in\mathcal{N}} g_{\check{k},n}$;
    Set $\rho_{\check{k},\hat{n}_{\check{k}}} \leftarrow 1$; $\mathcal{S}_{\check{k}} \leftarrow \mathcal{S}_{\check{k}} \cup \{\hat{n}_{\check{k}}\}$; $\mathcal{N} \leftarrow \mathcal{N} \setminus \{\hat{n}_{\check{k}}\}$;
    if $R_{\check{k}}\left(\mathcal{S}_{\check{k}}, P_{\check{k}}^{\max}\right) < \check{R}_{\check{k}}$
      $\eta_{EE,\check{k}}^{UL} \leftarrow \frac{R_{\check{k}}\left(\mathcal{S}_{\check{k}},P_{\check{k}}^{\max}\right)}{\zeta_{\check{k}} P_{\check{k}}^{\max}+P_{\check{k}}^c}$;
    else
      $\eta_{EE,\check{k}}^{UL} \leftarrow \max_{R_{\check{k}}^{-1}(\mathcal{S}_{\check{k}},\check{R}_{\check{k}})\le P_{\check{k}}\le P_{\check{k}}^{max}} \frac{R_{\check{k}}(\mathcal{S}_{\check{k}},P_{\check{k}})}{\zeta_{\check{k}} P_{\check{k}}+P_{\check{k}}^c}$;
      $\mathcal{F} \leftarrow \mathcal{F} \setminus \{\check{k}\}$;
    end
  else
    Find the UE $\check{k} \in \mathcal{K}$ such that $\check{k} \leftarrow \arg\min_{k\in\mathcal{K}} \eta_{EE,k}^{UL}$;
    Find the subcarrier $\hat{n}_{\check{k}} \in \mathcal{N}$ such that $\hat{n}_{\check{k}} \leftarrow \arg\max_{n\in\mathcal{N}} g_{\check{k},n}$;
    Set $\rho_{\check{k},\hat{n}_{\check{k}}} \leftarrow 1$; $\mathcal{S}_{\check{k}} \leftarrow \mathcal{S}_{\check{k}} \cup \{\hat{n}_{\check{k}}\}$; $\mathcal{N} \leftarrow \mathcal{N} \setminus \{\hat{n}_{\check{k}}\}$;
    Calculate $\eta_{EE,\check{k}}^{UL} \leftarrow \max_{R_{\check{k}}^{-1}(\mathcal{S}_{\check{k}},\check{R}_{\check{k}})\le P_{\check{k}}\le P_{\check{k}}^{\max}} \left\{\frac{R_{\check{k}}(\mathcal{S}_{\check{k}},P_{\check{k}})}{\zeta_{\check{k}} P_{\check{k}}+P_{\check{k}}^c}\right\}$;
  end
end

Table 3.5. Complexity Comparison for Uplink Transmission.

Algorithm	Complexity
optimal: brute-force search based on BPA	$\mathcal{O}(N_{OL}K^{N+1})$
suboptimal: MUSA	$\mathcal{O}(N_{OL}N)$

3.4 Numerical Results

In this section, we present simulation results to verify benefit of the energy-efficient design and performance of the low-complexity algorithms. In our simulation, the frequency spacing between adjacent subcarriers is 15 kHz. The frequency-selective Rayleigh fading of the kth UE is modeled with the ITU Pedestrian-B model [75] with an average CNR of $\bar{g}_k$. The total bandwidth, 1.08 MHz, is equally divided into 72 orthogonal subcarriers. For the downlink transmission, there are four UEs each with the same minimum rate requirement of 100 kbps. The first two UEs are with the same average CNR while the other two UE are with the same 10 dB higher average CNR.[4] The circuit power is 10 W and the maximum transmit power is 20 W for the base station. For the uplink transmission, there are four UEs each with the same average CNR and minimum rate requirement of 50 kbps. The circuit power is 50 mW and the maximum transmit power is 150 mW for each UE. For simplicity, we assume that the drain efficiency of power amplifier is 38% [61] for both the downlink and uplink cases.

Figures 3.1 and 3.2 evaluate the EE of the energy-efficient design that optimizes the generalized EE and the spectral-efficient design that maximizes the WSR with the same constraints expect for the objectives in the downlink transmission. From them, the energy-efficient design significantly improves EE compared to the spectral-efficient design. And the suboptimal energy-efficient scheme based on the MDSA and the BPA algorithms results in an EE that is at least 90% of the optimal EE. In Figure 3.2 (showing the case with the unequal UE weights), we also plot the resultant actual/conventional EE corresponding to the generalized EE of each scheme and find that the actual EE of the EE-suboptimal scheme

[4]From the following Figure 3.1 to Figure 3.5, the average CNR in the horizontal axis represents the CNR of the two low CNR UEs.

is also close to that of the EE-optimal one.

Figure 3.3 plots the throughput corresponding to the EE in Figures 3.1 and 3.2. From it, the throughput of the energy-efficient design is, as expected, less than that of the spectral-efficient design. The EE-suboptimal scheme has a slightly larger throughput than the EE-optimal one.[5] Together with Figures 3.1 and 3.2, we find that the maximum EE and the maximum throughput (or SE) are not necessarily simultaneously achieved, which indicates a tradeoff relationship between EE and SE as investigated in [79].

Figures 3.4 and 3.5 compare the average individual rate of the low CNR UEs and high CNR UEs corresponding to EE in Figures 3.1 and 3.2. From Figure 3.4, in such a case with equal weights, the rate of low CNR UEs is almost fixed at the minimum rate requirement, 100 kbps, while the rate of high CNR UEs increases with the CNR, which is unfair to the low CNR UEs. However, by adjusting the UE weights, as shown in Figure 3.5, the rate of low CNR UEs may also increase with the CNR and even become larger than that of the high CNR UEs. Hence, by properly choosing the UE weights rather than setting them equal for the energy-efficient design, certain fairness/priority notions can be introduced while still reflecting the aspects of the conventional EE.

Figures 3.6 and 3.7 illustrate the EE and throughput of the spectral-efficient design in [71] and the energy-efficient design based on the suboptimal MUSA algorithm, respectively. From them, the performance difference in EE and minimum individual rate between the energy-efficient design and the spectral-efficient design increase with the CNR, which also indicates the tradeoff relation between EE and SE for the uplink transmission.

3.5 Conclusion

In this chapter, we have studied the energy-efficient resource allocation in both downlink and uplink OFDMA networks. For each scenario, we first find the optimal energy-efficient

[5] Compared with the EE-optimal approach, the throughput improvement of the EE-suboptimal one is at the cost of more transmit power.

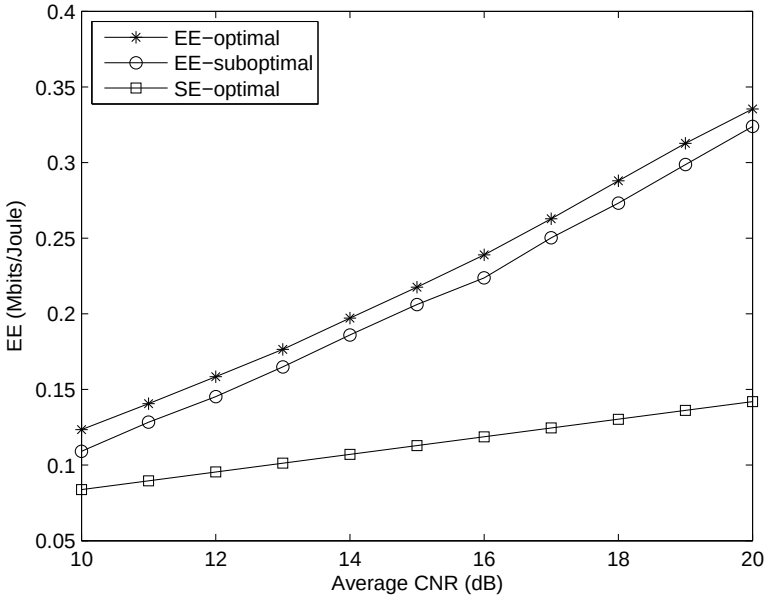


Figure 3.1. Comparison of the EE for the downlink transmission schemes with $\omega_1 = \omega_2 = \omega_3 = \omega_4 = 1$ and $\bar{g}_1 = \bar{g}_2 = \frac{1}{10}\bar{g}_3 = \frac{1}{10}\bar{g}_4$.

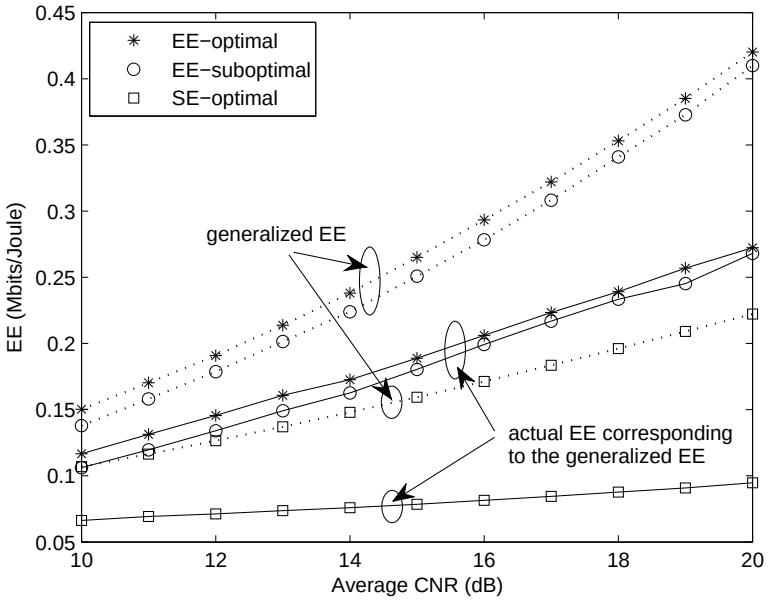


Figure 3.2. Comparison of the EE for the downlink transmission schemes with $\omega_1 = \omega_2 = 2.5$, $\omega_3 = \omega_4 = 1$, and $\bar{g}_1 = \bar{g}_2 = \frac{1}{10}\bar{g}_3 = \frac{1}{10}\bar{g}_4$.

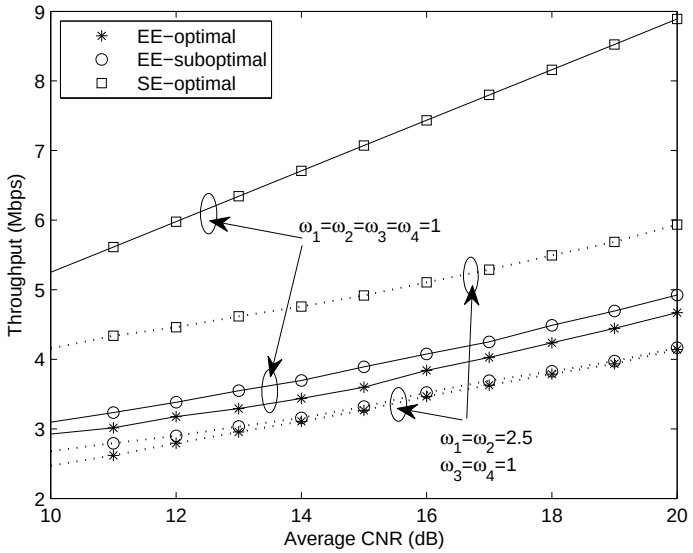


Figure 3.3. Comparison of the throughput for the downlink transmission schemes with $\bar{g}_1 = \bar{g}_2 = \frac{1}{10}\bar{g}_3 = \frac{1}{10}\bar{g}_4$.

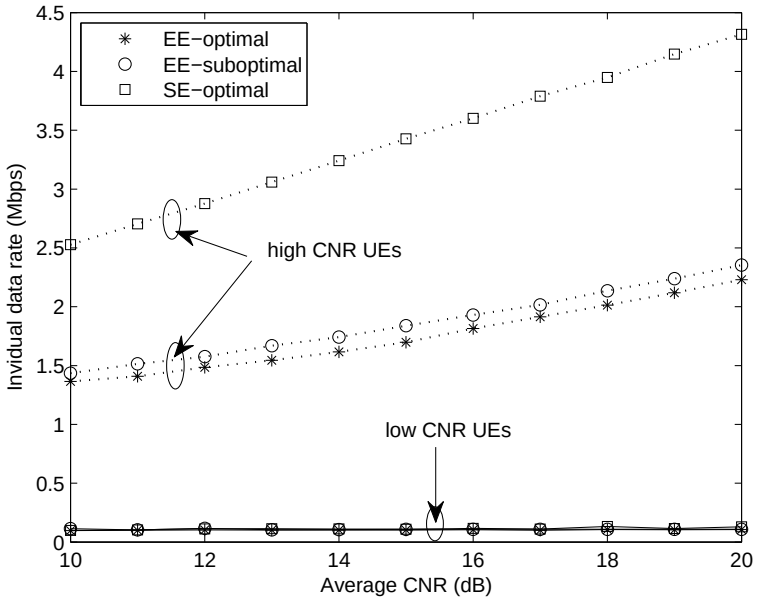


Figure 3.4. Comparison of the individual data rate for the downlink transmission schemes with $\omega_1 = \omega_2 = \omega_3 = \omega_4 = 1$ and $\bar{g}_1 = \bar{g}_2 = \frac{1}{10}\bar{g}_3 = \frac{1}{10}\bar{g}_4$.

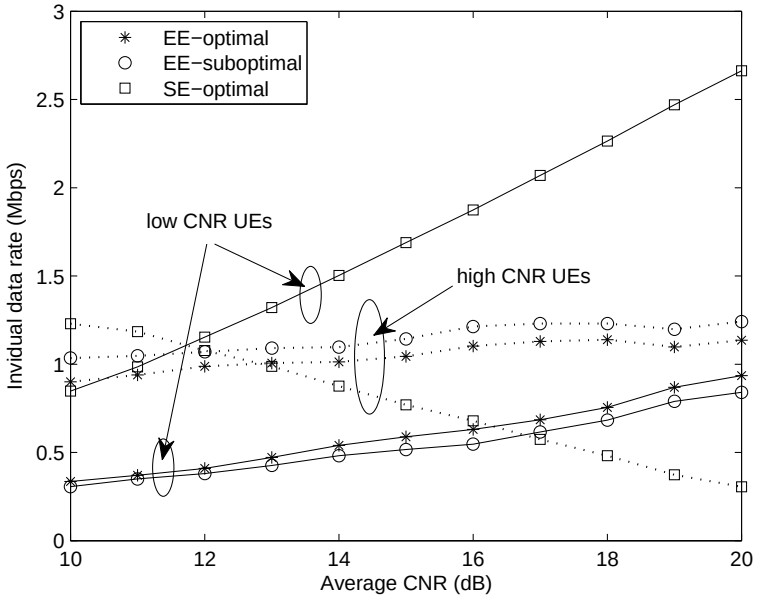


Figure 3.5. Comparison of the individual data rate for the downlink transmission schemes with $\omega_1 = \omega_2 = 2.5$, $\omega_3 = \omega_4 = 1$, and $\bar{g}_1 = \bar{g}_2 = \frac{1}{10}\bar{g}_3 = \frac{1}{10}\bar{g}_4$.

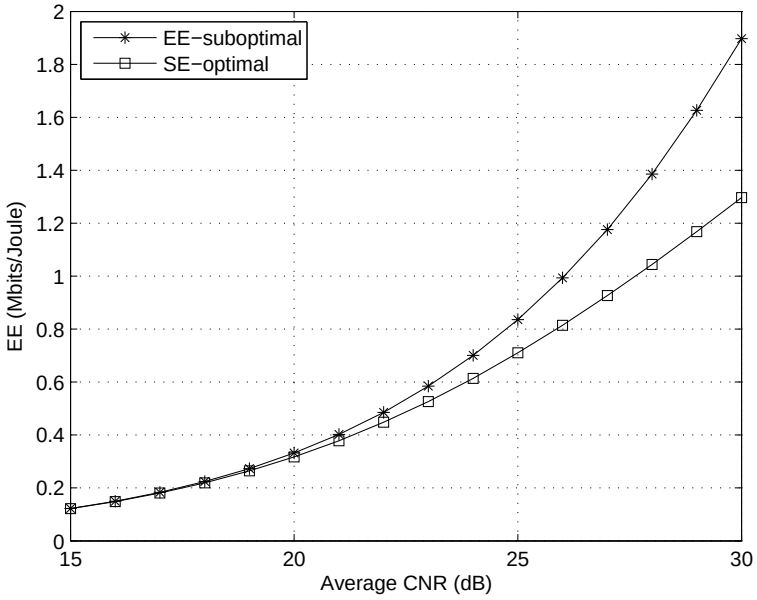


Figure 3.6. Comparison of the EE for the uplink transmission schemes with $\bar{g}_1 = \bar{g}_2 = \bar{g}_3 = \bar{g}_4$.

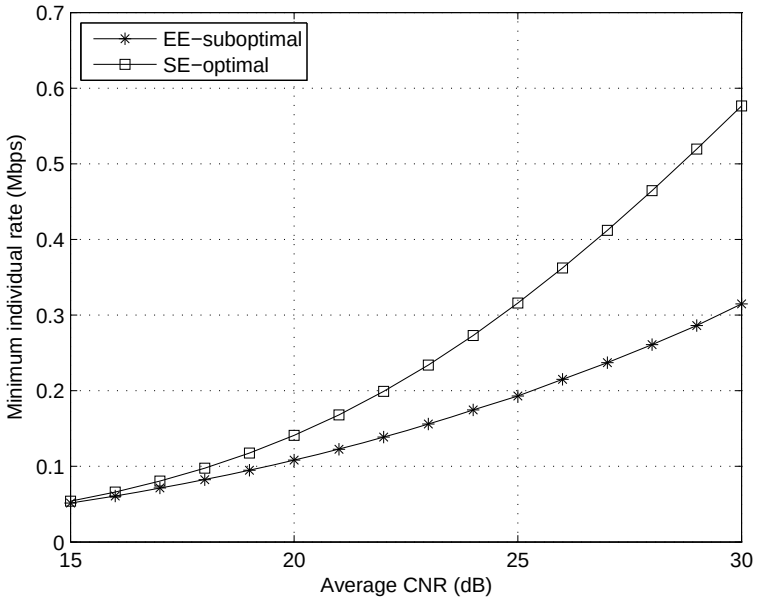


Figure 3.7. Comparison of the minimum individual data rate for the uplink transmission schemes with $\bar{g}_1 = \bar{g}_2 = \bar{g}_3 = \bar{g}_4$.

resource allocation approach then develop low-complexity suboptimal algorithm by exploring the inherent structure of the objective function and the feature of energy-efficient design. For the downlink case, we also obtain a computationally efficient and numerically tractable EE upper bound, which tends to be rather tight if the number subcarriers is large compared with that of UEs and is the foundation of a near-optimal solution relying on the quasiconcave relation between the modified EE and transmit power. Simulation results show great EE improvement of the energy-efficient design than that of the spectral-efficient design and the proposed low-complexity algorithms can achieve a promising tradeoff between performance and complexity.

CHAPTER 4

ENERGY-EFFICIENT DESIGN FOR DOWNLINK OFDMA WITH DELAY-SENSITIVE TRAFFIC

In this chapter, we address optimal energy-efficient resource allocation in a downlink network with the consideration of statistical delay guarantees for delay-sensitive traffic and also include its spectral-efficient design counterpart for comparison and connection. We model the problems as maximizing effective capacity-based throughput and EE under statistical delay constraints. We then solve the problems by tackling tractable upper bounds on the original problems based on continuous and concave relaxations. The intrinsic quasiconcavity of EE on transmit power is proved and exploited for solution, which is further demonstrated to be relatively tight when the number of subcarriers is larger than that of users. We also study the fundamental tradeoff relation between EE and delay, the connection between spectral-efficient and energy-efficient designs, and the impact of system parameters including circuit power and delay exponents.

The rest of this chapter is organized as follows. In Section 4.1, we describe the system model and formulate the optimization problem. In Section 4.2, we briefly investigate optimal power and subcarrier allocation for statistical delay-guaranteed spectral-efficient design and reveal the SE-delay tradeoff relation. In Section 4.3, we study optimal power and subcarrier allocation for statistical delay-guaranteed energy-efficient design as well as analyze EE-delay tradeoff. In section 4.4, we address the relationship between spectral-efficient and energy-efficient designs and the impact of system parameters. Then, we present numerical results in Section 4.5 and conclude this chapter in Section 4.6.

4.1 Problem Description

In this section, we introduce a system model of downlink OFDMA networks and formulate the problem of energy-efficient design with statistical delay provisioning.

4.1.1 System Model

We consider a single cell downlink OFDMA system consisting of K active users with heterogeneous services. The total bandwidth W is equally divided into N subcarriers, each with a bandwidth of $B = W/N$. We assume that the channel power gain of each user on each subcarrier is a stationary, ergodic, and uncorrelated discrete-time process and perfect *channel state information* (CSI) is available at the receiver while only the probability distribution of the channel power gain are known at the transmitter for the sake of reducing CSI feedback overhead of a bunch of subcarriers. For better tractability and simplicity, we also adopt the frequent assumption in the literature that the channel power gains of all subcarriers of the same user are independent and ignores the correlation when developing our scheme.

Assume that each subcarrier is exclusively assigned to at most one user in each time frame to avoid interference among different users. Denote instantaneous transmit power and channel power gain of the kth user on the nth subcarrier as $p_{k,n}$ and $g_{k,n}$, respectively. Then, the maximum instantaneous service/transmission rate of the kth user on the nth subcarrier in one frame with duration T is accordingly

$$r_{k,n} = TB \log_2 \left(1 + \frac{g_{k,n} p_{k,n}}{N_0 B}\right) \text{ (bits/frame)}, \tag{4.1}$$

where N_0 is the single-sided noise spectral density. Although the exact value of $g_{k,n}$ is not known at the transmitter, rateless codes such as *Luby transform* (LT) and Raptor codes enable the transmitter to adapt its rate to the channel realization and achieve $r_{k,n}$ without requiring CSI at the transmitter [47]. For systems without such codes, service rates are smaller than $r_{k,n}$, and the results in this chapter can serve as an upper bound on SE and EE, respectively.

Then, the aggregate service rate for the kth user is $R_k = \sum_{n=1}^{N} \phi_{k,n} r_{k,n}$, where $\phi_{k,n} \in \{1, 0\}$ indicates whether the nth subcarrier is assigned to the kth user or not. Obviously, a feasible

subcarrier assignment indicator matrix, $\boldsymbol{\phi} = [\phi_{k,n}]_{K\times N}$, should satisfy

$$\boldsymbol{\phi} \in \Phi \triangleq \Big\{[\phi_{k,n}]_{K\times N} | \phi_{k,n} \in \{0,1\}, \sum_{k=1}^{K} \phi_{k,n} \leq 1\Big\}. \tag{4.2}$$

Denote $\theta_k > 0$ to be the statistical delay exponent of the kth user. The delay exponent θ_k characterizes the steady-state delay violation probability of the kth user by

$$\Pr(D_k \geq D_k^{\max}) \approx e^{-\theta_k c_k D_k^{\max}}, \tag{4.3}$$

where D_k represents the delay and is a random variable, $D_k^{\max}$ is the delay bound, and c_k is a constant determined by the arrival and service processes [80].[1] Furthermore, in the case with a constant arrival rate, the steady-state delay violation probability satisfies [81]

$$\Pr(D_k \geq D_k^{\max}) \leq \bar{c}_k e^{-\frac{1}{2}\theta_k a_k D_k^{\max}}, \tag{4.4}$$

where a_k is the arrival rate and $\bar{c}_k$ is a certain positive constant. Clearly, a larger θ implies a more stringent delay guarantee and vice versa. Then, as discussed in Appendix A, the effective capacity, which is the maximum constant arrival rate that the service can support under the statistical delay requirement specified by θ_k, of the kth user, is as follows:

$$\mathrm{E}_{\mathrm{C}}^{k}(\theta_k) \triangleq -\frac{1}{\theta_k} \ln \mathbb{E}\left(e^{-\theta_k \sum_{n=1}^{N} \phi_{k,n} r_{k,n}}\right) = -\frac{1}{\theta_k} \sum_{n=1}^{N} \ln \mathbb{E}(e^{-\theta_k \phi_{k,n} r_{k,n}}) \quad , \tag{4.5}$$

where $\mathbb{E}(\cdot)$ is the expectation operator. Let $\gamma_{k,n} \triangleq \frac{g_{k,n}}{N_0 B}$ be the *channel gain-to-noise ratio* (CNR) of the kth user on the nth subcarrier and $f_{k,n}(\gamma)$ be the probability density distribution of $\gamma_{k,n}$. We assume that $f_{k,n}(\gamma)$ is continuously differentiable in γ, which is true for almost all practical situations. Therefore, the expectation in (4.5) with respect to the random variables $\gamma_{k,n}$'s can be easily evaluated with the following expression:

$$\mathrm{E}_{\mathrm{C}}^{k}(\theta_k) = -\frac{1}{\theta_k} \sum_{n=1}^{N} \ln\left[\int_0^{\infty} e^{-\theta_k \phi_{k,n} r_{k,n}} f_{k,n}(\gamma_{k,n}) \mathrm{d}\gamma_{k,n}\right]. \tag{4.6}$$

[1] From [80], c_k equals the rate where the effective bandwidth function, $\mathrm{E}_{\mathrm{B}}(\theta)$, and effective capacity function, $\mathrm{E}_{\mathrm{C}}(\theta)$, intersect each other, i.e., $c_k = \mathrm{E}_{\mathrm{B}}(\theta^*) = \mathrm{E}_{\mathrm{C}}(\theta^*)$.

In practical networks, the total transmit power, P, of a base station is both nonnegative and bounded. Thus, any possible power allocation matrix, $\boldsymbol{P} = [p_{k,n}]_{K\times N}$, should be subject to

$$\boldsymbol{P} \in \mathcal{P} \triangleq \left\{ [p_{k,n}]_{K\times N} | p_{k,n} \geq 0, P = \sum_{k=1}^{K}\sum_{n=1}^{N} p_{k,n} \leq P_{\max} \right\}, \tag{4.7}$$

where $P_{\max}$ represents the maximum total transmit power at the transmitter.

Besides transmit power, circuit energy consumption is also included, which is incurred by active circuit blocks, such as analog-to-digital converter, digital-to-analog converter, synthesizer, and mixer [24]. The overall power consumption at the base station is given by [24, 3][2]

$$P_{tot} = \rho P + P_c = \rho \sum_{k=1}^{K} P_k + P_c, \tag{4.8}$$

where $1/\rho$ is the drain efficiency of power amplifier, $P_k = \sum_{n=1}^{N} p_{k,n}$ is the transmit power for the kth user, and P_c represents the circuit power consumption.

4.1.2 Problem Formulation

For the downlink OFDMA network, the overall effective capacity for a given delay exponent vector $\boldsymbol{\theta} \triangleq [\theta_k]_{K\times 1}$ is $\mathrm{E}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P}) \triangleq \sum_{k=1}^{K} \mathrm{E}_{\mathrm{C}}^{k}(\theta_k)$. Then, the spectral-efficient design can be established as maximizing the overall effective capacity under the individual statistical delay constraints

$$\widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})} \triangleq \max_{\boldsymbol{\phi}\in\Phi, \boldsymbol{P}\in\mathcal{P}} \mathrm{E}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P}) \quad \text{(bits/frame)}, \tag{4.9a}$$

subject to

$$-\frac{1}{\theta_k}\sum_{n=1}^{N} \ln \mathbb{E}(e^{-\theta_k \phi_{k,n} r_{k,n}}) \geq A_k, \ \forall k, \tag{4.9b}$$

where A_k is service arrival rate for the kth user. Compared to the conventional spectral-efficient design that normally maximizes the overall throughput/transmission rate [**?**, 69, 82,

[2]From [5], circuit power may change the shape of the EE-versus-rate curve. Thus, it should be modeled and taken into account when optimizing EE.

?], this one considers the delay requirement specified by $\boldsymbol{\theta}$ and maximizes the throughput subject to delay requirement.

On the other hand, throughout this chapter the EE for the downlink OFDMA network is defined as the ratio of the overall effective capacity to the totally consumed energy as follows

$$\eta_{EE}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P}) \triangleq \frac{\mathrm{E}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P})}{TP_{tot}} = \frac{\sum_{k=1}^{K} -\frac{1}{\theta_k} \sum_{n=1}^{N} \ln \mathbb{E}(e^{-\theta_k \phi_{k,n} r_{k,n}})}{T(\rho P + P_c)} \text{ (bits/Joule)} \quad . \tag{4.10}$$

Compared to the conventional definition of EE as the ratio of the overall throughput/transmission rate to the totally consumed energy [5, 83, 3], this definition considers the delay requirement specified by $\boldsymbol{\theta}$ and serves as a delay-guaranteed EE metric. Then, the energy-efficient design can be formulated into maximizing the EE under the statistical delay guarantees as follows

$$\widetilde{\eta}_{EE}^{(\boldsymbol{\theta})} \triangleq \max_{\boldsymbol{\phi} \in \Phi, \boldsymbol{P} \in \mathcal{P}} \eta_{EE}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P}), \tag{4.11a}$$

subject to

$$-\frac{1}{\theta_k} \sum_{n=1}^{N} \ln \mathbb{E}(e^{-\theta_k \phi_{k,n} r_{k,n}}) \geq A_k, \ \forall k. \tag{4.11b}$$

Problems in (4.9) and (4.11) can be simplified if the channel power gain of different subcarriers belonging to one user are also identically distributed random processes, which is true in many practical applications. In this case, $g_{k,n}$'s for the same user are *independent and identically distributed* (i.i.d) and we can simply use $f_k(\gamma)$ to uniformly denote $f_{k,n}(\gamma)$, $n = 1, 2, \cdots, N$. Then, equal power allocation among the occupied subcarriers should be applied for each user for both the spectral-efficient and energy-efficient designs.[3] Let N_k be the number of occupied subcarriers of the kth user. The power on each occupied subcarrier of the kth user is P_k/N_k. Then problem (4.9) can be simplified by substituting (4.5) into it

[3]For any concave function $f(x)$ and variables x_1 and x_2 such that $x_1 + x_2 = y$, we have $f(x_1) + f(x_2) \leq 2f(\frac{x_1+x_2}{2}) = 2f(y/2)$, where equality holds if $x_1 = x_2 = y/2$. And $-\ln \mathbb{E}(e^{-\theta_k \phi_{k,n} r_{k,n}})$ is just a concave function of $p_{n,k}$ as proved in Lemma 1. Hence, the equal power allocation policy holds in the i.i.d. case.

as

$$\widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})} \triangleq \max_{N_k,P_k} \sum_{k=1}^{K} -\frac{N_k}{\theta_k} \ln \mathbb{E}\left(e^{-\theta_k TB \log_2\left(1+\frac{P_k\gamma_k}{N_k}\right)}\right), \tag{4.12a}$$

subject to

$$-\frac{N_k}{\theta_k} \ln \mathbb{E}\left(e^{-\theta_k TB \log_2\left(1+\frac{P_k\gamma_k}{N_k}\right)}\right) \geq A_k, \ \forall k, \tag{4.12b}$$

$$\sum_{k=1}^{K} N_k = N, \ N_k \geq 0, \ N_k \in \mathbb{Z}, \tag{4.12c}$$

$$\sum_{k=1}^{K} P_k \leq P_{\max}, \ P_k \geq 0. \tag{4.12d}$$

Problem (4.11) can be simplified into

$$\widehat{\eta}_{EE}^{(\boldsymbol{\theta})} \triangleq \max_{N_k,P_k} \frac{\sum_{k=1}^{K} -\frac{N_k}{\theta_k} \ln \mathbb{E}\left(e^{-\theta_k TB \log_2\left(1+\frac{P_k\gamma_k}{N_k}\right)}\right)}{T\left(\rho \sum_{k=1}^{K} P_k + P_c\right)}, \tag{4.13a}$$

subject to

$$-\frac{N_k}{\theta_k} \ln \mathbb{E}\left(e^{-\theta_k TB \log_2\left(1+\frac{P_k\gamma_k}{N_k}\right)}\right) \geq A_k, \ \forall k, \tag{4.13b}$$

$$\sum_{k=1}^{K} N_k = N, \ N_k \geq 0, \ N_k \in \mathbb{Z}, \tag{4.13c}$$

$$\sum_{k=1}^{K} P_k \leq P_{\max}, \ P_k \geq 0. \tag{4.13d}$$

The channel gain for the kth user, γ_k, is subject to the distribution of $f_k(\gamma)$. From (4.9) to (4.12) and (4.11) to (4.13), the number of optimization variables are reduced from $2NK$ to $2K$.

Mathematically, the problems in (4.9), (4.12), (4.11), and (4.13) may be infeasible if any of the service arrival rate is positive. In the infeasible cases, the limited transmit power and subcarrier resources are just not sufficient to support all users' delay requirement simultaneously. For simplicity and with practical consideration, we just assume that these problems are always feasible throughout this chapter as a result of other procedures, e.g., admission control.

4.2 Delay-Guaranteed Spectral-Efficient Design

In this section, we first introduce a method for achieving the optimal power and subcarrier allocation for delay-guaranteed spectral-efficient design in (4.9) and (4.12), which will be helpful to the energy-efficient design. Then, we analyze the SE-delay tradeoff relation.

4.2.1 Optimal Power and Subcarrier Allocation

To achieve spectral-efficient design, we need to solve problem (4.9) to find the optimal power and subcarrier allocation, which is a combinatorial integer programming problem and is in general NP-hard. To make it more tractable, continuous relaxation of $\phi_{k,n}$ and concave relaxation of $\mathrm{E}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P})$ to be jointly concave in $\boldsymbol{\phi}$ and $\boldsymbol{P}$ are helpful. Motivated by relaxations in literature such as [70, 71], we similarly modify (4.2), (4.1), and $\mathrm{E}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P})$ into

$$\tilde{\boldsymbol{\phi}} \in \tilde{\Phi} \triangleq \left\{ [\tilde{\phi}_{k,n}]_{K\times N} | \tilde{\phi}_{k,n} \in [0,1], \sum_{k=1}^{K} \tilde{\phi}_{k,n} \leq 1 \right\}, \tag{4.14}$$

$$\tilde{r}_{k,n} = TB \log_2\left(1 + \frac{\gamma_{k,n} p_{k,n}}{\tilde{\phi}_{k,n} N_0 B}\right), \tag{4.15}$$

$$\tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right) \triangleq \sum_{k=1}^{K} -\frac{1}{\theta_k} \sum_{n=1}^{N} \ln \mathbb{E}(e^{-\theta_k \tilde{\phi}_{k,n} \tilde{r}_{k,n}}), \tag{4.16}$$

respectively. After the above relaxations, the optimization problem turns into

$$\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})} \triangleq \max_{\tilde{\boldsymbol{\phi}} \in \tilde{\Phi}, \boldsymbol{P} \in \mathcal{P}} \tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right), \tag{4.17a}$$

subject to

$$-\frac{1}{\theta_k} \sum_{n=1}^{N} \ln \mathbb{E}(e^{-\theta_k \tilde{\phi}_{k,n} \tilde{r}_{k,n}}) \geq A_k, \ \forall k. \tag{4.17b}$$

It is obvious that the following problem (4.17) yields an upper bound on the maximum overall effective capacity of (4.9), i.e., $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})} \geq \widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}$.

From the following Lemma 4.1 that is proved in Appendix C.2, problem (4.17) is in the standard form of a convex programming that can be solved by techniques such as the interior-point method [76].

Lemma 4.1 *The modified effective capacity,* $\tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right)$*, in (4.16) is jointly concave in* $\tilde{\phi}_{k,n}$*'s and* $p_{k,n}$*'s.*

The optimal $\tilde{\phi}_{k,n}^{opt}$'s to obtain $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}$ are not ensured to be integers. To get a feasible solution to problem (4.9), we need to round the possibly fractional $\tilde{\phi}_{k,n}^{opt}$'s to 0 or 1. Such manipulations may not result in the optimal solution to (4.9). Luckily, this is rarely a problem if the number of subcarriers is large compared with the number of users because the optimal subcarrier allocation indicators, $\tilde{\phi}_{k,n}^{opt}$'s, almost tend to be either 0 or 1 when $K \ll N$ [70, 71]. In such a case, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}$ is quite close to $\widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}$.

Similarly, by relaxing N_k's from integers to fractional numbers, problem (4.12) turns out to be a convex programming and can be preciously solved by techniques such as the interior-point method.

4.2.2 SE and Delay Tradeoff Relation

Since $\mathrm{E}_{\mathrm{C}}^{k}(\theta_k)$ is a strictly decreasing function of θ_k [84, 46], the following SE-delay tradeoff relation can be easily proved.

Property 4.1 *The maximum overall effective capacity,* $\widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}$*, in (4.9) (and (4.12)) is a strictly and monotonically decreasing function of all delay exponents,* θ_k*'s.*

From Property 4.1, we cannot simultaneously improve SE as well as reduce delay, which agrees with our intuition. If one increases, the other will decrease, and vice versa. And the SE can be maximized when no delay guarantee is required.

4.3 Delay-Guaranteed Energy-Efficient Design

In this section, we first present the fundamental framework for achieving the optimal power and subcarrier allocation for delay-guaranteed energy-efficient design in (4.11) and (4.13). Then, we analyze the EE-delay tradeoff relation, the relationship between the energy-efficient and spectral-efficient designs, and the impact of system parameters.

4.3.1 Framework for Optimal Power and Subcarrier Allocation

To achieve energy-efficient design, we need to solve problem (4.11) to find the optimal power and subcarrier allocation, which is a fractional programming problem and is in general NP-hard. To make it more tractable, we utilize the same modifications as in the spectral-efficient design as follows:

$$\tilde{\eta}_{EE}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right) \triangleq \frac{\tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right)}{T\left(\rho P + P_c\right)}. \tag{4.18}$$

Then, a modified version of (4.11) can be established as follows

$$\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})} \triangleq \max_{\tilde{\boldsymbol{\phi}} \in \tilde{\Phi}, \boldsymbol{P} \in \mathcal{P}} \tilde{\eta}_{EE}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right), \tag{4.19a}$$

subject to

$$-\frac{1}{\theta_k} \sum_{n=1}^{N} \ln \mathbb{E}(e^{-\theta_k \tilde{\phi}_{k,n} \tilde{r}_{k,n}}) \geq A_k, \ \forall k. \tag{4.19b}$$

Obviously, due to the relaxation of $\phi_{k,n}$ to $\tilde{\phi}_{k,n}$ and the modification of $r_{k,n}$ into $\tilde{r}_{k,n}$, problem (4.19) results in an upper bound on the EE of (4.13), i.e, $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})} \geq \widehat{\eta}_{EE}^{(\boldsymbol{\theta})}$. For preparation of solving (4.19), we first define $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{P})$, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$, and $\widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$ as follows

$$\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{P}) \triangleq \max_{\tilde{\boldsymbol{\phi}} \in \tilde{\Phi}} \tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right), \tag{4.20}$$

under constraint (4.19b). And

$$\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P) \triangleq \max_{\tilde{\boldsymbol{\phi}} \in \tilde{\Phi}, p_{k,n} \geq 0} \tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}, \boldsymbol{P}\right), \tag{4.21}$$

under constraints (4.19b) and $\sum_{k=1}^{K} \sum_{n=1}^{N} p_{k,n} \leq P$. And

$$\widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P) \triangleq \max_{\boldsymbol{\phi} \in \Phi, p_{k,n} \geq 0} \mathrm{E}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P}), \tag{4.22}$$

under constraints (4.11b) and $\sum_{k=1}^{K} \sum_{n=1}^{N} p_{k,n} \leq P$. Then, we have the following lemma, which is proved in Appendix C.3.

Lemma 4.2 *The maximum modified effective capacity, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{P})$, for a given transmit power matrix, $\boldsymbol{P}$, is strictly concave in $\boldsymbol{P}$. Furthermore, the maximum modified effective capacity, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$, for a given total transmit power, P, is strictly concave in P.*

Based on Lemma 4.2, we have the following theorem, which reveals the intrinsic property of (4.19) and is the foundation of the proposed optimal solution framework. It is proved in Appendix .

Theorem 4.1 *The upper bound on the maximum EE, $\widehat{\eta}_{EE}^{(\boldsymbol{\theta})}(P) \triangleq \widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)/\left(T\left(\rho P + P_c\right)\right)$, for a given total transmit power, P, namely,*

$$\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P) \triangleq \frac{\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{T\left(\rho P + P_c\right)}, \tag{4.23}$$

has the following properties

(i) $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$ is continuously differentiable and quasiconcave in P,

(ii) $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$ either decreases or first increases and then decreases with P,

(iii) $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}$
$$\begin{cases} > 0 & \text{if } \widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P) < \frac{1}{\rho T}\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} \\ = 0 & \text{if } \widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P) = \frac{1}{\rho T}\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} \\ < 0 & \text{if } \widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P) > \frac{1}{\rho T}\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} \end{cases},$$

and the derivative satisfies (4.24),

$$\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} = \max_{k,n}\left\{\tilde{\phi}_{k,n}^{*} TB\log_2 e \frac{\int_0^{\infty} e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_k}{\tilde{\phi}_{k,n}^{*}}\right)}\frac{\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}+p_{k,n}^{*}\gamma_{k,n}} f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}{\int_0^{\infty} e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_k}{\tilde{\phi}_{k,n}^{*}}\right)} f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}\right\}. \tag{4.24}$$

where $\tilde{\phi}_{k,n}^{*}$'s and $p_{k,n}^{*}$'s are the optimal power and subcarrier allocation of (4.23), respectively.

Theorem 4.1 demonstrates the quasiconcavity of $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$, which is an upper bound on the EE $\widehat{\eta}_{EE}^{(\boldsymbol{\theta})}(P)$, in the transmit power, P, and implies the existence and the uniqueness of the global maximum, i.e., $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}$, for (4.19). More importantly, as a result of the quasiconcavity, problem (4.19) can be decomposed into two layers and solved iteratively by the *joint inner- and outer-layer optimization* (JIOO) algorithm similar to [85] as follows:

(i) Inner layer: For a given transmit power, $P \leq P_{\max}$, find the maximum EE, $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$, and (the sign of) its derivative, $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}$.

(ii) Outer layer: Search for the transmit power, P_{opt}, that results in the maximum, $\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta})}$, by derivative-aided bisection power search.

The derivative-aided bisection search in the outer-layer is clear and easy according to (iii) in Theorem 4.1 once the inner-layer problem is solved. The key lies in the inner-layer algorithm that finds $\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$ and (the sign of) $\frac{\mathrm{d}\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}$. For a given total transmit power, the inner-layer problem is equivalent to maximizing the modified effective capacity in (4.21). As proved in Lemma 4.2, it is in a form of a convex programming problem and thus can be successfully solved by techniques such as the interior-point method. When the optimal subcarrier assignment matrix, $\tilde{\boldsymbol{\phi}}^*$, and power allocation matrix, $\boldsymbol{P}^*$, for (4.21) are obtained, the sign of the first derivative, $\frac{\mathrm{d}\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}$, can be readily determined from (iii) in Theorem 4.1. Hence, problem (4.19) can be accordingly solved by the aforementioned JIOO algorithm, which is summarized in Table 4.1.

Like the spectral-efficient design case, the optimal $\tilde{\phi}_{k,n}^{opt}$'s to obtain $\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta})}$ are not necessarily integers. For a feasible solution to the original downlink EE maximization problem (4.11), we have to round the possibly fractional $\tilde{\phi}_{k,n}^{opt}$'s to 0 or 1. Such manipulations may not lead to the optimal solution to (4.11). Luckily, due to the good tightness of $\widehat{\widetilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$ on $\widehat{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$ [70, 71], this is rarely a problem if the number of subcarriers is large compared with the number of users (because the optimal subcarrier allocation indicators, $\tilde{\phi}_{k,n}^{opt}$'s, almost tend to be either 0 or 1) and in such a case $\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta})}$ is quite close to $\widehat{\eta}_{EE}^{(\boldsymbol{\theta})}$.

Similarly, we can prove that problem (4.13) also fit the JIOO algorithm and can be well solved if we relax N_k to fractional numbers $\tilde{N}_k$. It is easy to derive the new expression for the derivative in (4.24) as in (4.25),

$$\frac{\mathrm{d}\widehat{\widetilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} = \max_k \left\{ \tilde{N}_k^* TB \log_2 \frac{\int_0^\infty e^{-\theta_k TB \log_2\left(1+\frac{P_k^*\gamma_k}{\tilde{N}_k^*}\right)} \frac{\gamma_k}{\tilde{N}_k^*+P_k^*\gamma_k} f_k(\gamma_k)\mathrm{d}\gamma_k}{\int_0^\infty e^{-\theta_k TB \log_2\left(1+\frac{P_k^*\gamma_k}{\tilde{N}_k^*}\right)} f_k(\gamma_k)\mathrm{d}\gamma_k} \right\}. \tag{4.25}$$

where $\tilde{N}_k^*$ is the optimal solution to obtain the corresponding $\widehat{\widetilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$.

Table 4.1. JIOO Algorithm.

Algorithm *JIOO*

1. Solve problem (4.21) or (4.23) with $P = P_{\max}$ to get the corresponding subcarrier and power allocation matrices $\tilde{\boldsymbol{\phi}}^*_{\max}$ and $\boldsymbol{P}^*_{\max}$; calculate $\widehat{\tilde{\eta}}^{(\boldsymbol{\theta})}_{EE}(P_{\max})$ and $\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}^{(\boldsymbol{\theta})}_{\mathrm{C}}(P_{\max})}{\mathrm{d}P}$ according to (4.23) and (4.24), respectively.
2. **if** $\frac{\mathrm{d}\widehat{\tilde{\eta}}^{(\boldsymbol{\theta})}_{EE}(P_{\max})}{\mathrm{d}P} \geq 0$
3. **then return** $\tilde{\boldsymbol{\phi}}^*_{\max}$ and $\boldsymbol{P}^*_{\max}$. ($*$ maximum EE is achieved $*$)
4. **else** $P^{(1)} \leftarrow P_{\max}$, $\tilde{\boldsymbol{\phi}}^{(1)} \leftarrow \tilde{\boldsymbol{\phi}}^*_{\max}$, and $\boldsymbol{P}^{(1)} \leftarrow \boldsymbol{P}^*_{\max}$.
5. Solve the transmit power minimization (convex optimization) problem, $P_{\min} \triangleq \min_{\tilde{\boldsymbol{\phi}}\in\Phi, p_{k,n}\geq 0} \sum_{k=1}^{K}\sum_{n=1}^{N} p_{k,n}$ under constraint (4.19b), to get $P_{\min}$ and the corresponding matrices $\tilde{\boldsymbol{\phi}}^*_{\min}$ and $\boldsymbol{P}^*_{\min}$; calculate $\widehat{\tilde{\eta}}^{(\boldsymbol{\theta})}_{EE}(P_{\min})$ and $\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}^{(\boldsymbol{\theta})}_{\mathrm{C}}(P_{\min})}{\mathrm{d}P}$.
6. **if** $\frac{\mathrm{d}\widehat{\tilde{\eta}}^{(\boldsymbol{\theta})}_{EE}(P_{\min})}{\mathrm{d}P} \leq 0$ (from (iii) in Theorem 4.1)
7. **then return** $\tilde{\boldsymbol{\phi}}^*_{\min}$ and $\boldsymbol{P}^*_{\min}$. ($*$ maximum EE is achieved $*$)
8. **else** $P^{(2)} \leftarrow P_{\min}$, $\tilde{\boldsymbol{\phi}}^{(2)} \leftarrow \tilde{\boldsymbol{\phi}}^*_{\min}$, and $\boldsymbol{P}^{(2)} \leftarrow \boldsymbol{P}^*_{\min}$.
9. **while** no convergence
10. **do** $P \leftarrow \frac{P^{(1)}+P^{(2)}}{2}$; solve problem (4.21) or (4.23) to get the corresponding $\tilde{\boldsymbol{\phi}}^*$ and $\boldsymbol{P}^*$; calculate $\widehat{\tilde{\eta}}^{(\boldsymbol{\theta})}_{EE}(P)$ and $\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}^{(\boldsymbol{\theta})}_{\mathrm{C}}(P)}{\mathrm{d}P}$.
11. **if** $\frac{\mathrm{d}\widehat{\tilde{\eta}}^{(\boldsymbol{\theta})}_{EE}(P)}{\mathrm{d}P} < 0$
12. **then** $P^{(1)} \leftarrow P$, $\tilde{\boldsymbol{\phi}}^{(1)} \leftarrow \tilde{\boldsymbol{\phi}}^*$, and $\boldsymbol{P}^{(1)} \leftarrow \boldsymbol{P}^*$.
13. **else** $P^{(2)} \leftarrow P$, $\tilde{\boldsymbol{\phi}}^{(2)} \leftarrow \tilde{\boldsymbol{\phi}}^*$, and $\boldsymbol{P}^{(2)} \leftarrow \boldsymbol{P}^*$.
14. **return** $\tilde{\boldsymbol{\phi}}^*$ and $\boldsymbol{P}^*$.

4.3.2 EE and Delay Tradeoff Relation

As discussed in Appendix C.5, the EE-delay tradeoff can be stated in the following property.

Property 4.2 *The maximum EE, $\widehat{\tilde{\eta}}^{(\boldsymbol{\theta})}_{EE}$, in (4.11) (and (4.13)) is a strictly and monotonically decreasing function of all delay exponents, θ_k's.*

From Property 4.2, we cannot simultaneously improve EE as well as reduce delay, which agrees with our intuition. If one increases, the other will decrease, and vice versa. And the EE can be maximized when no delay guarantee is required.

4.4 Relationship between Spectral-Efficient and Energy-Efficient Designs

From the analysis in Sections III and IV, the spectral-efficient design always transmits at the maximum transmit power regardless of the channel condition while the energy-efficient design adapts its total transmit power level according to the channel condition. And the solution of the spectral-efficient design can be viewed as a special case of the JIOO algorithm for the energy-efficient design, which only does the inner-layer optimization once with $P = P_{\max}$ and does not implements the outer-layer power search. Thus, the complexity of the energy-efficient design is basically that of the spectral-efficient design multiplying the number of outer-layer searches.

From the perspective of system performance in EE and overall effective capacity, spectral-efficient and energy-efficient designs have the following relationship:

- Overall effective capacity is a concave function of total transmit power and thus spectral efficient design always transmits at the maximum transmit power; EE is a quasiconcave function of total transmit power and thus energy-efficient design generally transmits at some power between the minimum power to ensure delay requirement and the maximum transmit power (not necessarily at the maximum power or at the minimum power).

- Spectral-efficient design can be viewed as a special case of energy-efficient design, where circuit power tends to be infinite or relatively large compared with the maximum transmit power. In such a case, the transmit power is negligible compared with the circuit power; thus maximizing SE is just equivalent to maximizing EE.

- A larger delay exponent will lead to a higher operating point of total transmit power (to ensure more stringent delay requirement) for energy-efficient design. Thus, energy-efficient design tends to be close to spectral-efficient design in EE and effective capacity with a large delay exponent.

- A larger circuit power will result in a higher operating point of total transmit power for energy-efficient design. Thus, energy-efficient design tends to be close to spectral-efficient design in EE and effective capacity with a high circuit power. Mathematically, this statement is summarized in Property 3 and proved in Appendix C.6.

Property 4.3 *The total transmit power, P_{opt}, for achieving the maximum EE, $\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta})}$, in (4.19) is a nondecreasing function of the circuit power, P_c.*

4.5 Numerical Results

In this section, we present numerical results to verify the benefit of the energy-efficient design and compare it with the spectral-efficient design. In the system configuration, the frequency spacing between adjacent subcarriers is 15 kHz. The frequency-selective fading is assumed to be with Rayleigh distribution and the subcarriers of the same user are assumed to be subject to i.i.d. frequency-selective fading. The total bandwidth, 0.6 MHz, is equally divided into 40 non-overlapping subcarriers and the frame duration is 1 ms. There are four users each with the same average CNR, delay exponent, and minimum effective capacity requirement of 50 kbps. The circuit power is 5 W and the maximum transmit power is 20 W for the base station. The drain efficiency of power amplifier is assumed to be 30%.

Figure 4.1 plots the maximum EE with respective to the total transmit power for both the aforementioned independent channel and another two types of correlated channel. From it, the EE is indeed a quasiconcave function of transmit power. It can be seen that the maximum EE is not necessarily achieved at the maximum total transmit power. On the other hand, the solution by rounding off the possible fractional number of subcarriers obtained from the solution based on the continuous and concave relaxations fits well with the upper bound. This demonstrates that the continuous and concave relaxations do result in a quite tight upper bound and can be used for optimal solution. For the correlated channel cases,

we plot the curves for the ITU Pedestrian-B channel (correlated channel) [75] and another extreme case, i.e., identical Rayleigh channel (flat-fading channel). It can be inferred that correlation of channel does not change the quasiconcave relationship between EE and transmit power but degrades the EE.

Figures 4.2 and 4.3 evaluate the EE and overall effective capacity of the energy-efficient and spectral-efficient designs with different delay exponents but with a fixed CNR. From them, the energy-efficient design significantly improves EE compared to the spectral-efficient design at the cost of some effective capacity degradation. From the figures, when the delay exponent is relatively small, the EE and the overall effective capacity almost remains unchanged with the delay exponent. This is because when the delay exponent is small, the delay requirement is loose and the overall effective capacity is close to Shannon capacity, which depends on the wireless channel but is independent of the arrival process. However, when the delay exponent becomes large, the EE and overall effective capacity both decrease much. And when the delay exponent is relatively large, the overall effective capacity of the energy-efficient design merely satisfies the minimum effective capacity requirement and remains unchanged, which can be explained by Property (ii) in Theorem 4.1 and indicates that EE decreases with the transmit power from the beginning.

Figures 4.4 and 4.5 compare the EE and the overall effective capacity of the energy-efficient design and the spectral-efficient design with different CNRs. We can also conclude that the energy-efficient design greatly outperforms the spectral-efficient design in EE and the large delay exponents reduce the EE and the overall effective capacity.

4.6 Conclusion

In this chapter, we have studied the statistical delay driven energy-efficient design in downlink OFDMA networks for delay-sensitive traffic and compared it with the spectral-efficient design. We formulate SE and EE optimization problems with the statistical delay provisioning based on the concept of the effective capacity. To solve the established problems,

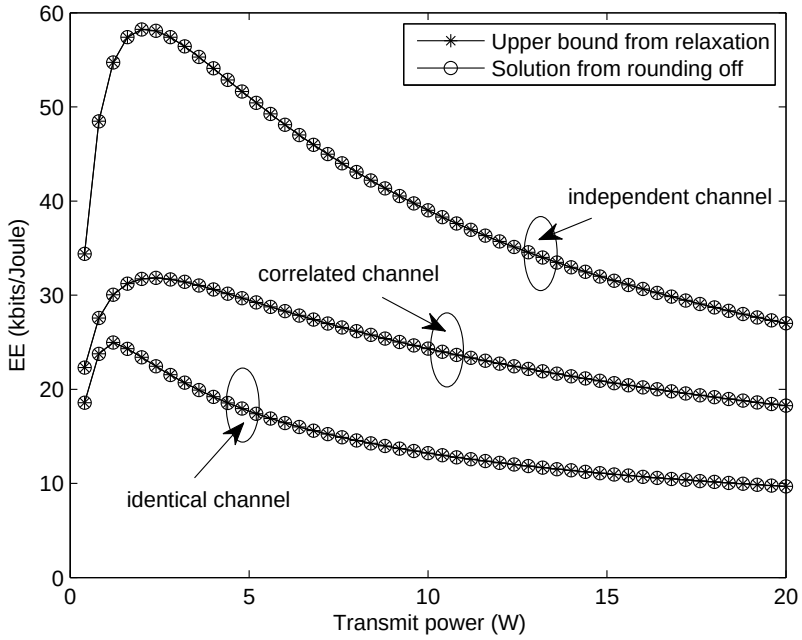


Figure 4.1. EE versus total transmit power with γ_k = 15 dB and θ_k = 0.1.

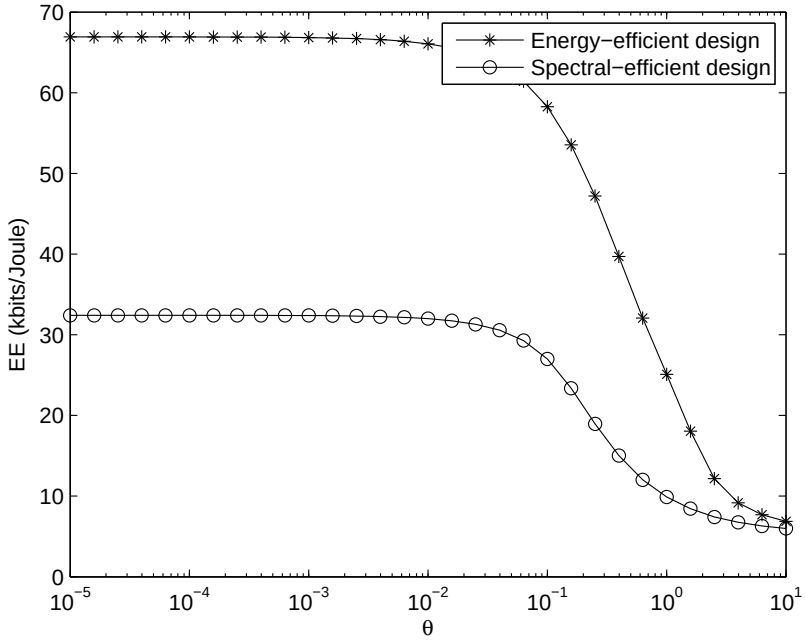


Figure 4.2. EE versus the delay exponent with γ_k = 15 dB.

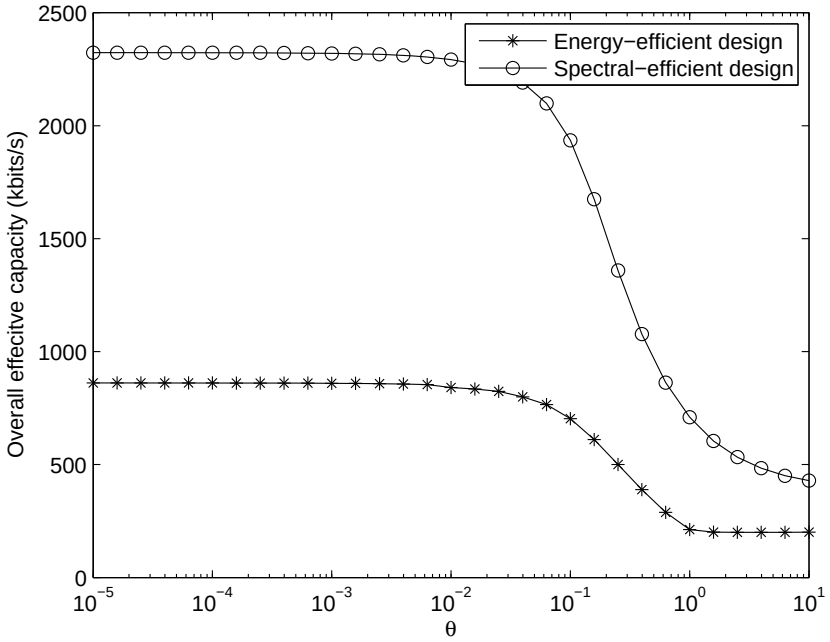


Figure 4.3. Overall effective capacity versus the delay exponent with γ_k = 15 dB.

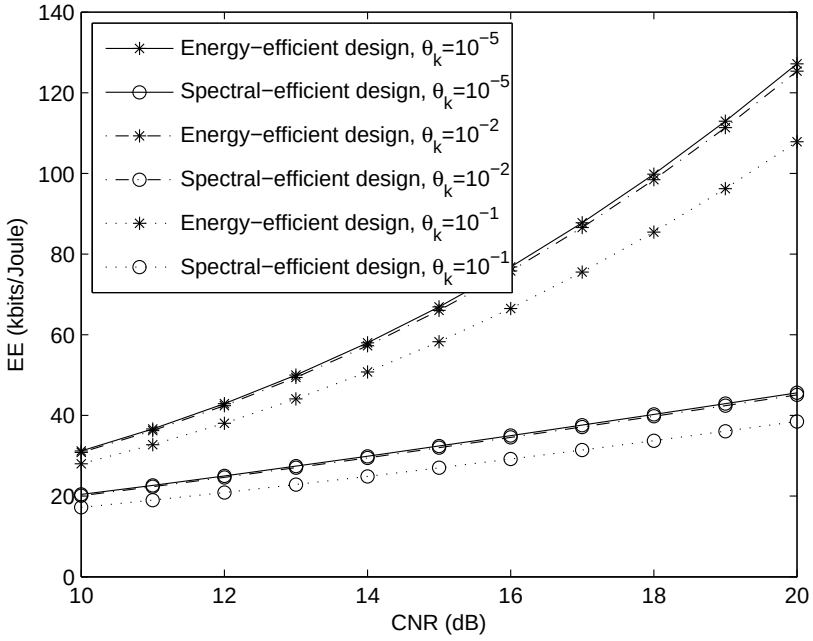


Figure 4.4. EE versus the the average CNR.

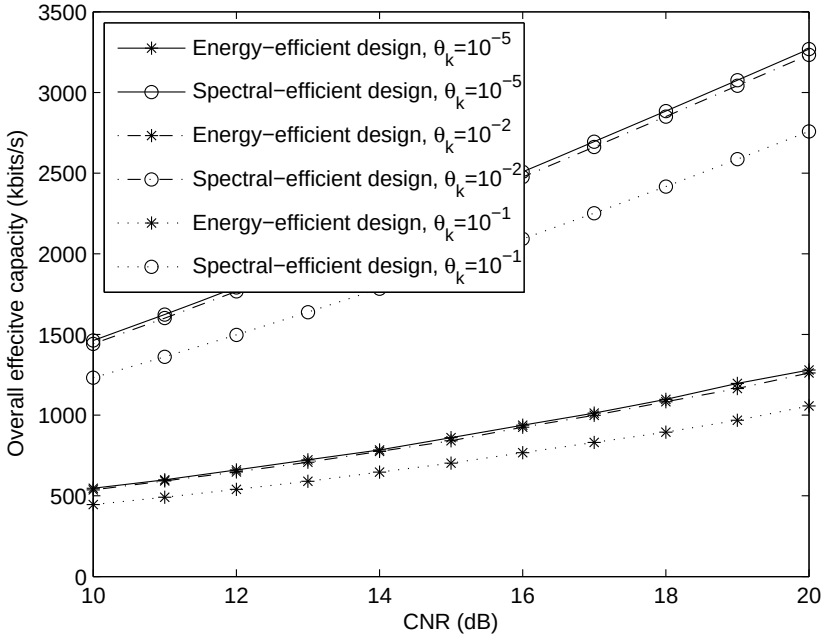


Figure 4.5. Overall effective capacity versus the average CNR.

we resort to continuous and concave relaxations that yield an upper bound on the original problems. Then, we prove and exploit the intrinsic quasiconcavity of EE on transmit power, which implicates the existence of a unique global maximum of EE. We demonstrate that the resultant upper bound is quite tight when the number of subcarriers is larger than that of the users. We also analyze the EE-delay tradeoff, the relationship between spectral-efficient and energy-efficient designs, and the impact of system parameters, such as circuit power and delay exponents. Numerical results show that the proposed energy-efficient design scheme significantly enhances EE with QoS guarantees but at some possible loss of throughput.

CHAPTER 5

ENERGY-EFFICIENT SPECTRUM ACCESS IN COGNITIVE RADIOS

In this chapter, we address spectrum access strategies for multiple SUs in OFDM-based CR, aiming at optimizing the worst EE and the average EE, respectively. As most of single-transmitter EE problems for cellular networks, they are nonconvex integer combinatorial fractional optimization issues. However, as a result of different mathematical structure, the optimization procedures for single-transmitter cellular networks cannot be directly used in CR networks. For better mathematical tractability, we utilize continuous and convex relaxation to modify the problems. For the relaxed worst-EE-based spectrum access problem, we first prove the joint quasiconcavity of EE on subchannel and power allocation matrices, which guarantees the global optimization of the established problem, and then develop a framework to find the optimal solution based on efficient root finding and convex optimization. We also develop a low-complexity alternative for suboptimal solution. The relaxed average-EE-based spectrum access problem is still NP-hard because of the sum-of-ratios structure and may have many local optima. We first convert it into an equivalent form and introduce a general concave envelope based *branch-and-bound* (B&B) approach to find the global optimal solution. We also exploit the underlying properties of the energy-efficient transmission to speed up the convergence of the B&B approach. In addition, we develop a low-complexity heuristic approach for a suboptimal solution.

The rest of the chapter is organized as follows. In Section 5.1, we describe the network model, introduce the definition of EE, and formulate two EE optimization problems and their modifications. In Sections 5.2 and 5.3, the near-optimal and low-complexity suboptimal solutions are investigated for the two problems, respectively. Then, we present numerical results in Section 5.4. Finally, we conclude this chapter in Section 5.5.

5.1 Problem Description

In this section, we first introduce the network model and the definition of EE. Then we formulate the problems of energy-efficient spectrum access and modify them for better mathematical tractability.

5.1.1 Network Description

We consider an OFDM-based CR network, underlaying another OFDM-based primary network. The K SUs in the CR network have their respective intended receivers or share a common receiver. Denote the index set of all SUs as $\mathcal{K} = \{1, 2, \cdots, K\}$. The total licensed bandwidth is equally divided into non-overlapping subchannels, each with a bandwidth of W. To protect the PU, the CR network accesses a licensed subchannel for opportunistic transmission only when it is not currently used by the primary network. To decide the presence of the primary signal or the availability for opportunistic secondary access on a subchannel, spectrum sensing [11, 12, 13, 14] needs to be performed by the CR network. Denote the index set of all subchannels available for secondary access as $\mathcal{N} \triangleq \{1, 2, \cdots, N\}$.

Each available subchannel is exclusively assigned to at most one SU at a time to avoid interference among different SUs. Each SU can simultaneously transmit on multiple subchannels. Let $\theta_{n,k} \in \{1, 0\}$ be the subchannel allocation indicator and notifies whether subchannel $n \in \mathcal{N}$ is assigned to SU by 1 or not by 0. Clearly, a feasible subchannel assignment indictor matrix, $\mathbf{\Theta} \triangleq (\theta_{n,k})_{n,k=1}^{N,K}$, should satisfy

$$\begin{aligned} \mathbf{\Theta} \in \Theta \triangleq \Big\{(\theta_{n,k})_{n,k=1}^{N,K} | \sum_{k\in\mathcal{K}} \theta_{n,k} \leq 1, \forall\, n \in \mathcal{N}; \\ \theta_{n,k} \in \{0, 1\}, \forall\, k \in \mathcal{K}, n \in \mathcal{N}\Big\}. \end{aligned} \tag{5.1}$$

The subchannel assignment constraint in (5.1) can also be equivalently interpreted as follows:

$$\bigcup_{k=1}^{K} \mathcal{S}_k \subseteq \mathcal{N} \text{ and } \mathcal{S}_k \bigcap \mathcal{S}_{k'} = \varnothing, \forall\, k \neq k',$$

where $\mathcal{S}_k$ is the set of subchannels assigned to SU k.

Denote the transmit power of SU k on subchannel n as $p_{n,k}$. For each subchannel $n \in \mathcal{N}$, at most one SU $k \in \mathcal{K}$ may have a positive transmit power, $p_{n,k} > 0$, on it while $p_{n,k'} = 0$ for $k' \neq k$. The total transmit power of SU k is

$$p_k = \sum_{n \in \mathcal{S}_k} p_{n,k} = \sum_{n \in \mathcal{N}} p_{n,k},$$

which cannot exceed the maximum overall transmit power allowed for each SU, $p_{\max}$.

If there is mis-detection, that is, the PU is using a subchannel while spectrum sensing ignores the primary signal, CR transmission in this case will interfere with the primary signal on that subchannel. To protect the primary transmission in this case, we impose a per subchannel based average interference power constraint for the CR network as follows:

$$\sum_{k \in \mathcal{K}} p_{n,k} \bar{g}_k \leq \bar{I}_{th}, \ \forall\, n \in \mathcal{N}, \tag{5.2}$$

where $\bar{g}_k$ is the average channel power gain from SU k to the primary receiver, $\bar{I}_{th}$ is the average maximum tolerable interference level at the active primary receiver on each subchannel. From (5.1), at most one term in the summation of (5.2) will be of a positive value. Hence, Eq. (5.2) basically limits the maximum power that a SU is allowed to transmit on each available subchannel. Therefore, we have

$$p_{n,k} \leq \frac{\bar{I}_{th}}{\bar{g}_k} \triangleq \delta_k.$$

Then, any possible power allocation matrix, $\mathbf{P} \triangleq (p_{n,k})_{n,k=1}^{N,K}$, should be subject to[1]

$$\mathbf{P} \in \mathcal{P} \triangleq \left\{ (p_{n,k})_{n,k=1}^{N,K} \;\middle|\; [p_{1,k}, \cdots, p_{N,k}]^T \in \mathcal{P}_k, \forall k \in \mathcal{K}; \ \sum_{k \in \mathcal{K}} p_{n,k} \bar{g}_k \leq \bar{I}_{th}, \ \forall\, n \in \mathcal{N} \right\}, \tag{5.3}$$

where $\mathcal{P}_k \triangleq \left\{ (p_{n,k})_{n=1}^{N} \;\middle|\; 0 \leq p_{n,k} \leq \delta_k, \forall n \in \mathcal{N}; \ p_k \leq p_{\max} \right\}$.

The average data rate of SU k on subchannel $n \in \mathcal{N}$ is accordingly

$$r_{n,k} = \theta_{n,k} q W \log_2 \left(1 + p_{n,k} \gamma_{n,k}\right), \tag{5.4}$$

[1]The constraint (5.2) in $\mathcal{P}$ is redundant here and for problems (5.6) and (5.7), but will be necessary for the relaxed problems (5.11) and (5.12).

where $1-q$ is the posteriori probability of mis-detection. $\gamma_{n,k} \triangleq \frac{g_{n,k}}{\sigma^2}$ and $g_{n,k} > 0$ are the received *channel-gain-to-noise ratio* (CNR) and the channel power gain of SU k on subchannel n, respectively. σ^2 is the noise power on each subchannel at the secondary receiver side. Perfect channel state information of each SU, $(\gamma_{n,k})_{n=1}^{N}$, is assumed to be accurately known to the CR network throughout this chapter. Consequently, the average aggregate data rate of SU k is

$$r_k = \sum_{n \in \mathcal{S}_k} r_{n,k} = \sum_{n \in \mathcal{N}} r_{n,k}.$$

5.1.2 EE Metrics and Problem Formulation

As in [26, 85], the EE of SU k, $\mathcal{E}_k$, is defined as follows:

$$\begin{aligned} \mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k) &\triangleq \frac{r_k}{p_c + \rho p_k} \\ &= \frac{\sum_{n \in \mathcal{N}} qW\theta_{n,k} \log_2(1 + p_{n,k}\gamma_{n,k})}{p_c + \rho \sum_{n \in \mathcal{N}} p_{n,k}}, \end{aligned} \tag{5.5}$$

where p_c and $1/\rho$ are the circuit power and amplifier efficiency of the SU transmitter, respectively. $\boldsymbol{\theta}_k \triangleq [\theta_{1,k}, \theta_{2,k}, \cdots, \theta_{N,k}]^T$ and $\mathbf{p}_k \triangleq [p_{1,k}, p_{2,k} \cdots, p_{N,k}]^T$.

In some application scenarios, we may prefer certain level of absolute fairness with respect to EE among different SUs and thus guarantee the performance of the worst SU. The weighted worst EE of the CR network, $\underline{\mathcal{E}}$, is defined as

$$\underline{\mathcal{E}}(\boldsymbol{\Theta}, \mathbf{P}) \triangleq \min_{k \in \mathcal{K}} w_k \mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k),$$

where w_k is a predetermined positive weight for SU k. The basic idea of the worst-EE-based spectrum access is to maximize $\underline{\mathcal{E}}$ under the subchannel allocation constraint in (5.1) and the per subchannel interference and peak transmit power constraint in (5.3). Then, we have the following worst-EE optimization problem:

$$(\underline{\boldsymbol{\Theta}}_o, \underline{\mathbf{P}}_o) \triangleq \arg \max_{\boldsymbol{\Theta} \in \Theta, \mathbf{P} \in \mathcal{P}} \underline{\mathcal{E}}(\boldsymbol{\Theta}, \mathbf{P}), \tag{5.6}$$

where $\underline{\boldsymbol{\Theta}}_o$, $\underline{\mathbf{P}}_o$, and $\underline{\mathcal{E}}_o$ represent the optimal subchannel and power matrices and the optimal worst EE, respectively.

While maximizing the worst EE, $\underline{\mathcal{E}}$, can provide somewhat absolute fairness with regard to EE among different SUs, the average EE of the CR network is sometimes more important. The weighted average EE of the CR network, $\overline{\mathcal{E}}$, is expressed as

$$\overline{\mathcal{E}}(\mathbf{\Theta}, \mathbf{P}) \triangleq \sum_{k \in \mathcal{K}} w_k \mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k),$$

which naturally characterizes the overall EE level of all SUs. This definition is based on the summation of the weighted EE of all SUs rather than the ratio of the overall weighted rate of all SUs to the overall power of all SUs because the powers of different SUs in CR cannot be shared and neither can their rate and EE. Similar to (5.6), we can perform resource allocation to maximize the average EE as follows:

$$(\overline{\mathbf{\Theta}}_o, \overline{\mathbf{P}}_o) \triangleq \arg\max_{\mathbf{\Theta} \in \Theta, \mathbf{P} \in \mathcal{P}} \overline{\mathcal{E}}(\mathbf{\Theta}, \mathbf{P}), \tag{5.7}$$

where $\overline{\mathbf{\Theta}}_o$, $\overline{\mathbf{P}}_o$, and $\overline{\mathcal{E}}_o$ denote the optimal subchannel and power matrices and the optimal average EE, respectively.

5.1.3 Modifications for Better Mathematical Tractability

The worst-EE and the average-EE problems are both nonconvex integer combinatorial and NP-hard. Therefore, we first introduce the continuous and concave modification to provide better mathematical tractability and aid developing low-complexity but near-optimal approaches. The subchannel allocation indicator, $\theta_{n,k}$ is binary variable. As in [69, 70], we relax $(\theta_{n,k})_{n,k=1}^{N,K} \in \Theta$ into continuous ones, $(\varphi_{n,k})_{n,k=1}^{N,K} \in \Phi$, as

$$\mathbf{\Phi} \in \Phi \triangleq \left\{ (\varphi_{n,k})_{n,k=1}^{N,K} \middle| \sum_{k \in \mathcal{K}} \varphi_{n,k} \leq 1, \forall\, n \in \mathcal{N}; \varphi_{n,k} \in [0,1], \forall\, k \in \mathcal{K}, n \in \mathcal{N} \right\}. \tag{5.8}$$

The data rate in (5.4) can be modified accordingly as

$$\widetilde{r}_{n,k} = \varphi_{n,k} q W \log_2 \left(1 + \frac{p_{n,k} \gamma_{n,k}}{\varphi_{n,k}} \right). \tag{5.9}$$

The fractional $\varphi_{n,k}$ in (5.9) can be interpreted as frequency-domain sharing of a subchannel [70].

Similar to (5.5), the modified EE of SU k is defined as follows:

$$\xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k) \triangleq \frac{\sum_{n\in\mathcal{N}} qW\varphi_{n,k}\log_2\left(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}}\right)}{p_c+\rho\sum_{n\in\mathcal{N}}p_{n,k}}, \tag{5.10}$$

where $\boldsymbol{\varphi}_k \triangleq [\varphi_{1,k}, \varphi_{2,k}, \cdots, \varphi_{N,k}]^T \in \Phi_k$ and $\Phi_k \triangleq \left\{(\varphi_{n,k})_{n=1}^N \middle| \varphi_{n,k} \in [0,1], \forall\, n \in \mathcal{N}\right\}$. Then, the worst and the average modified EE of the CR network are respectively

$$\underline{\xi}(\boldsymbol{\Phi}, \mathbf{P}) \triangleq \min_{k\in\mathcal{K}} w_k\xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k) \text{ and } \overline{\xi}(\boldsymbol{\Phi}, \mathbf{P}) \triangleq \sum_{k\in\mathcal{K}} w_k\xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k).$$

Similar to (5.6), based on the modified EE, the relaxed worst-EE-guaranteed problem is

$$(\underline{\boldsymbol{\Phi}}_o, \underline{\widetilde{\mathbf{P}}}_o) \triangleq \arg\max_{\boldsymbol{\Phi}\in\Phi, \mathbf{P}\in\mathcal{P}} \underline{\xi}(\boldsymbol{\Phi}, \mathbf{P}), \tag{5.11}$$

where $\underline{\boldsymbol{\Phi}}_o$, $\underline{\widetilde{\mathbf{P}}}_o$, and $\underline{\xi}_o$ are the optimal subchannel and power allocation matrices and the resultant optimal worst modified EE, respectively. Similarly, the relaxed average-EE optimization problem is

$$(\overline{\boldsymbol{\Phi}}_o, \overline{\widetilde{\mathbf{P}}}_o) \triangleq \arg\max_{\boldsymbol{\Phi}\in\Phi, \mathbf{P}\in\mathcal{P}} \overline{\xi}(\boldsymbol{\Phi}, \mathbf{P}), \tag{5.12}$$

where $\overline{\boldsymbol{\Phi}}_o$, $\overline{\widetilde{\mathbf{P}}}_o$, and $\overline{\xi}_o$ are the optimal subchannel and power allocation matrices and the resultant optimal average modified EE, respectively.

Apparently, the relaxed problems will always result in an upper bound on the respective original problems, i.e., $\underline{\xi}_o \geq \underline{\mathcal{E}}_o$ and $\overline{\xi}_o \geq \overline{\mathcal{E}}_o$.

5.2 Worst-EE-Guaranteed Spectrum Access

In this section, we investigate the worst-EE-guaranteed spectrum access in (5.6) and (5.11). We first provide a near-optimal solution and then develop a low-complexity approach for a suboptimal solution.

5.2.1 Near-Optimal Solution

Theorem 5.1 *The worst modified EE, $\underline{\xi}(\boldsymbol{\Phi}, \mathbf{P})$, under constraints (5.3) and (5.8), is jointly quasiconcave in $\boldsymbol{\Phi}$ and $\mathbf{P}$ and has a global maximum, $\underline{\xi}_o$.*

Theorem 5.1, proved in Appendix D.1, demonstrates the existence of the global optimal of the relaxed worst-EE problem in (5.11). Moreover, it indicates that a local maximum is also globally optimal. The next theorem, proved in Appendix D.2, facilitates the optimal solution of (5.11).

Theorem 5.2 *Let* $J_k(\eta) \triangleq \sum_{n\in\mathcal{N}} w_k q W \varphi_{n,k} \log_2\left(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}}\right) - \eta\left(p_c + \rho \sum_{n\in\mathcal{N}} p_{n,k}\right)$ *and* $\eta \geq 0$*, then* $\underline{J}(\eta) \triangleq \min_{k\in\mathcal{K}} J_k(\eta)$ *under constraints (5.3) and (5.8), satisfies*

$$\widehat{\underline{J}}(\eta) \triangleq \max_{\mathbf{\Phi}\in\Phi,\mathbf{P}\in\mathcal{P}} \underline{J}(\eta) = 0 \text{ iff } \eta = \underline{\xi}_o.$$

Moreover, $\widehat{\underline{J}}(\eta)$ *strictly decreases with* η *and* $\underline{J}(\eta)$ *is jointly concave in* $\mathbf{\Phi}$ *and* $\mathbf{P}$*.*

From Theorem 5.2, if $\widehat{\underline{J}}(\eta) > 0$, then $\eta < \underline{\xi}_o$; if $\widehat{\underline{J}}(\eta) < 0$, then $\eta > \underline{\xi}_o$. Hence, $\underline{\xi}_o$ can be easily obtained by various root-finding methods to get the root of $\widehat{\underline{J}}(\eta) = 0$ as long as the value of $\widehat{\underline{J}}(\eta)$ is accessible for any $\eta \geq 0$. Since $\underline{J}(\eta)$ is jointly concave in $\mathbf{\Phi}$ and $\mathbf{P}$, $\widehat{\underline{J}}(\eta)$ can be optimally obtained via convex optimization techniques, such as the interior point method [76]. The procedure of solving the original worst-EE problem in (5.6) near optimally based on solving the relaxed worst-EE problem in (5.11) is summarized in Table 5.1.

5.2.2 Low-Complexity Suboptimal Solution

The complexity to optimally solve $\widehat{\underline{J}}(\eta)$ is prohibitively high in practice. To shed some light on the design of low-complexity approaches, we first study the case with a given channel assignment by introducing the next theorem, which can similarly proved as Theorem 5.1.

Theorem 5.3 *For any given subchannel allocation matrix,* $\mathbf{\Theta}_{\text{v}} \triangleq [\boldsymbol{\theta}_1^{\text{v}}, \boldsymbol{\theta}_2^{\text{v}}, \cdots, \boldsymbol{\theta}_K^{\text{v}}]$*, the individual EE,* $\mathcal{E}_k^{(\boldsymbol{\theta}_k^{\text{v}})}(\mathbf{p}_k) \triangleq \mathcal{E}_k(\boldsymbol{\theta}_k^{\text{v}}, \mathbf{p}_k)$*, is strictly quasiconcave in* $\mathbf{p}_k \in \mathcal{P}_k$ *and therefore has a unique global maximum,* $\mathcal{E}_{k,o}^{(\boldsymbol{\theta}_k^{\text{v}})} \triangleq \max_{\mathbf{p}_k\in\mathcal{P}_k} \mathcal{E}_k^{(\boldsymbol{\theta}_k^{\text{v}})}(\mathbf{p}_k)$*. The worst EE,* $\underline{\mathcal{E}}^{(\mathbf{\Theta}_{\text{v}})}(\mathbf{P}) \triangleq \min_{k\in\mathcal{K}} w_k \mathcal{E}_k^{(\boldsymbol{\theta}_k^{\text{v}})}(\mathbf{p}_k)$*, is quasiconcave in* $\mathbf{P} \in \mathcal{P}$ *and* $\mathbf{p}_k \in \mathcal{P}_k$ *and has a global maximum. Moreover, the maximum of the worst EE equals the minimum of the maximum weighted EE, i.e,* $\underline{\mathcal{E}}_o^{(\mathbf{\Theta}_{\text{v}})} \triangleq \max_{\mathbf{P}\in\mathcal{P}} \underline{\mathcal{E}}^{(\mathbf{\Theta}_{\text{v}})}(\mathbf{P}) = \min_{k\in\mathcal{K}} w_k \mathcal{E}_{k,o}^{(\boldsymbol{\theta}_k^{\text{v}})}$*.*

Table 5.1. Near-Optimal Solution for the Worst-EE-Based Spectrum Access.

1. For the chosen $\eta \geq 0$, solve $\widehat{\underline{J}}(\eta)$ by convex optimization algorithms and get the corresponding optimal subchannel and power allocation matrices, $\mathbf{\Phi}^*(\eta) = (\varphi^*_{n,k}(\eta))^{N,K}_{n,k=1}$ and $\mathbf{P}^*(\eta) = (p^*_{n,k}(\eta))^{N,K}_{n,k=1}$.
2. Update η with the worst modified EE, i.e., $\eta \leftarrow \min\limits_{k\in\mathcal{K}} \left\{ w_k \xi^*_k(\eta) \triangleq w_k \frac{\sum\limits_{n\in\mathcal{N}} qW\varphi^*_{n,k}(\eta)\log_2\left(1+\frac{p^*_{n,k}(\eta)\gamma_{n,k}}{\varphi^*_{n,k}(\eta)}\right)}{p_c+\rho\sum\limits_{n\in\mathcal{N}} p^*_{n,k}(\eta)} \right\}$.
3. Repeat Steps 1-2 till convergence is reached. After converging, $\mathbf{\Phi}^*(\underline{\xi}_o) = \underline{\mathbf{\Phi}}_o$ and $\mathbf{P}^*(\underline{\xi}_o) = \underline{\widetilde{\mathbf{P}}}_o$ are obtained, which lead to an upper bound on $\underline{\mathcal{E}}_o$.
4. Round off the subchannel allocation matrix, $\mathbf{\Phi}^*(\underline{\xi}_o)$, to a feasible subchannel allocation matrix $\mathbf{\Theta}^*(\underline{\xi}_o) \triangleq [\boldsymbol{\theta}^*_1(\underline{\xi}_o), \boldsymbol{\theta}^*_2(\underline{\xi}_o), \cdots, \boldsymbol{\theta}^*_K(\underline{\xi}_o)] \in \Theta$.
5. For each SU $k \in \mathcal{K}$, recalculate and update its transmit power vector with the obtained subchannel vector, $\boldsymbol{\theta}^*_k(\underline{\xi}_o)$, to reach its maximize individual EE under constraint (5.3), i.e., $\mathcal{E}_k^{(\boldsymbol{\theta}^*_k(\underline{\xi}_o))}$. Output the resultant subchannel and power allocation matrices, $\mathbf{\Theta}^*(\underline{\xi}_o)$ and $\mathbf{P}^*(\underline{\xi}_o)$, as the feasible near-optimal solution.

According to Theorem 5.3, the key of obtaining the maximum of the worst EE, $\underline{\mathcal{E}}_o^{(\mathbf{\Theta}_\mathrm{v})}$ lies in getting the maximum EE, $\mathcal{E}_{k,o}^{(\boldsymbol{\theta}^\mathrm{v}_k)}$, for all SUs. We define $\mathcal{E}_{k,o}^{(\mathcal{S}^\mathrm{v}_k)} \triangleq \mathcal{E}_{k,o}^{(\boldsymbol{\theta}^\mathrm{v}_k)}$, where $\mathcal{S}^\mathrm{v}_k$ is the designated subchannel set for SU k containing the subchannels with $\theta^\mathrm{v}_{k,n} = 1$ in $\boldsymbol{\theta}^\mathrm{v}_k$. To get $\mathcal{E}_{k,o}^{(\mathcal{S}^\mathrm{v}_k)}$, we resort to the next theorem that can be similarly proved as Theorem 5.2.

Theorem 5.4 *Let $\eta_k \geq 0$, then*

$$J_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k) \triangleq \sum_{n\in\mathcal{S}^\mathrm{v}_k} w_k qW\log_2(1+p_{n,k}\gamma_{n,k}) - \eta_k(p_c+\rho\sum_{n\in\mathcal{S}^\mathrm{v}_k} p_{n,k}),$$

under constraint (5.3), satisfies

$$\widehat{J}_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k) \triangleq \max_{\mathbf{p}_k\in\mathcal{P}_k} J_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k) = 0 \ \text{ iff } \ \eta_k = w_k\mathcal{E}_{k,o}^{(\mathcal{S}^\mathrm{v}_k)}.$$

Moreover, $\widehat{J}_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k)$ strictly decreases with η_k and $J_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k)$ is strictly concave in $\mathbf{p}_k$ and p_k.

From Theorem 5.4, if $\widehat{J}_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k) > 0$, then $\eta_k < w_k\mathcal{E}_{k,o}^{(\mathcal{S}^\mathrm{v}_k)}$; if $\widehat{J}_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k) < 0$, then $\eta_k > w_k\mathcal{E}_{k,o}^{(\mathcal{S}^\mathrm{v}_k)}$. Hence, $\mathcal{E}_{k,o}^{(\mathcal{S}^\mathrm{v}_k)}$ can be easily obtained by various root-finding methods, such as Newton's method, to get the root of $\widehat{J}_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k) = 0$ as long as the value of $\widehat{J}_k^{(\mathcal{S}^\mathrm{v}_k)}(\eta_k)$ is accessible for any

$\eta_k \geq 0$. Since $J_k^{(\mathcal{S}_k^{\mathrm{v}})}(\eta_k)$ is a concave function in p_k, $\widehat{J}_k^{(\mathcal{S}_k^{\mathrm{v}})}(\eta_k)$ is easy to obtain by adjusting the sum power $p_k \leq \min(p_{\max}, \sum_{n\in\mathcal{S}_k^{\mathrm{v}}} \delta_k) \triangleq \chi_k$. The procedure for determining the power level is as follows.

For a given power, $p_k \leq \chi_k$, we have the optimal power allocation and the corresponding derivative[2] for the function

$$\max_{\sum_{n\in\mathcal{S}_k^{\mathrm{v}}} p_{k,n}=p_k} J_k^{(\mathcal{S}_k^{\mathrm{v}})}(\eta_k)$$

under constraint (5.3) as follows:

$$\wp_{n,k} = \min\left(\delta_k, \left[\frac{w_k q W}{(\eta_k \rho + \lambda_k)\ln 2} - \frac{1}{\gamma_{n,k}}\right]^+\right), \tag{5.13}$$

$$\frac{\mathrm{d}J_k^{(\mathcal{S}_k^{\mathrm{v}})}(\eta_k)}{\mathrm{d}p_{n,k}} = \frac{w_k q W}{(\wp_{n,k} + 1/\gamma_{n,k})\ln 2} - \eta_k \rho, \tag{5.14}$$

where $\lambda_k \geq 0$ is the parameter (Lagrange multiplier) making $\sum_{n\in\mathcal{S}_k^{\mathrm{v}}} \wp_{n,k} = p_k$. Based on (5.13) and (5.14), we can branch the problem of getting $\widehat{J}_k^{(\mathcal{S}_k^{\mathrm{v}})}(\eta_k)$ into two cases and solve it accordingly as follows:

- If $\sum_{n\in\mathcal{S}_k^{\mathrm{v}}} \min\left(\delta_k, [\frac{w_k q W}{\eta_k \rho \ln 2} - \frac{1}{\gamma_{n,k}}]^+\right) \geq \chi_k$, the optimal sum transmit power is the maximum transmit power, i.e., $p_k = \chi_k$. And the optimal transmit power on each subcarrier is determined by water-filling in (5.13), i.e., $p_{n,k} = \wp_{n,k}$, where $\sum_{n\in\mathcal{S}_k^{\mathrm{v}}} \wp_{n,k} = \chi_k$.

- If $\sum_{n\in\mathcal{S}_k^{\mathrm{v}}} \min\left(\delta_k, [\frac{w_k q W}{\eta_k \rho \ln 2} - \frac{1}{\gamma_{n,k}}]^+\right) < \chi_k$, the optimal sum transmit power is strictly less than the maximum transmit power, i.e., $0 \leq p_k < \chi_k$. And the optimal transmit power on each subcarrier is $p_{k,n} = \min(\delta_k, [\frac{w_k q W}{\eta_k \rho \ln 2} - \frac{1}{\gamma_{n,k}}]^+)$. In particular, if $w_k q W \gamma_{n,k} \leq \eta_k \rho \ln 2$, $\forall\, n \in \mathcal{S}_k^{\mathrm{v}}$, the optimal transmit power on each subchannel is $p_{n,k} = 0$ and the sum transmit power is $p_k = 0$.

From the above analysis, it takes at most one water-filling to obtain $\widehat{J}_k^{(\mathcal{S}_k^{\mathrm{v}})}(\eta_k)$.

To facilitate the practical application of the worst-EE optimization in problems (5.6) and (5.11), we then propose a low-complexity suboptimal approach for getting $\underline{\widehat{J}}(\eta)$ and

[2]The derivatives of $J_k^{(\mathcal{S}_k^{\mathrm{v}})}(\eta_k)$ with respect to $p_{k,n}$ at $p_{k,n} = 0$ and $p_{k,n} = \delta_k$ are defined by the right-hand derivative and left-hand derivative, respectively.

Table 5.2. Low-Complexity Suboptimal Solution for the Worst-EE-based Spectrum Access.

1. For the chosen $\eta \geq 0$, set the subchannel set for each SU as an empty set, i.e., $\mathcal{S}_k \leftarrow \emptyset, \forall k \in \mathcal{K}$. Set the unassigned subchannel set as the set of all available subchannels, i.e., $\mathcal{N}_{us} \leftarrow \mathcal{N}$. For each SU $k \in \mathcal{K}$, estimate the value of its $\mathcal{J}_k(\eta)$ using $\widehat{J}_k^{(\mathcal{S}_k^{vt})}(\eta)$, i.e., $\mathcal{J}_k(\eta) \leftarrow \widehat{J}_k^{(\mathcal{S}_k^{vt})}(\eta)$, where $\mathcal{S}_k^{vt}$ is a virtual subchannel set containing only one subchannel with a channel power gain of $\gamma_k = \frac{1}{N}\sum_{n\in\mathcal{N}} \gamma_{n,k}$.
2. For the SU with the minimum $\mathcal{J}_k$, i.e., SU $k^* \leftarrow \text{argmin}_{k\in\mathcal{K}} \mathcal{J}_k$, assign it its best subchannel among all unassigned subchannels, i.e., $\mathcal{S}_{k^*} \leftarrow \mathcal{S}_{k^*} \cup \{n^*\}$, where $n^* \leftarrow \arg\max_{n\in\mathcal{N}_{us}} \gamma_{n,k^*}$. For SU k^*, recalculate its $\widehat{J}_{k^*}^{(\mathcal{S}_{k^*})}(\eta)$ and update $\mathcal{J}_{k^*}$ with it, i.e., $\mathcal{J}_{k^*} \leftarrow \widehat{J}_{k^*}^{(\mathcal{S}_{k^*})}(\eta)$. Remove subchannel n^* from the unassigned subchannel set, i.e. $\mathcal{N}_{us} \leftarrow \mathcal{N}_{us} \setminus \{n^*\}$.
3. Repeat Step 2 till all the subchannels are assigned, i.e., $\mathcal{N}_{us} = \emptyset$.
4. Update η with the resultant worst weighted EE, i.e., $\eta \leftarrow \min_{k\in\mathcal{K}} \left\{ w_k \mathcal{E}_k^*(\eta) \triangleq \frac{w_k \sum_{n\in\mathcal{S}_k} r_{n,k}}{p_c + \rho \sum_{k\in\mathcal{S}_k} p_{n,k}} \right\}$.
5. Repeat Steps 1-4 till convergence or the stopping criterion is satisfied.
6. For each SU $k \in \mathcal{K}$, recalculate and update its transmit power vector with the obtained subchannel set, $\mathcal{S}_k$, to reach its maximize EE under constraint (5.3), i.e., $\mathcal{E}_{k,o}^{(\mathcal{S}_k)}$. Output the resultant subchannel sets and power allocation matrix, $\{\mathcal{S}_k\}$ and $\mathbf{P}$, as the suboptimal solution.

its root. The basic idea of the proposed low-complexity algorithm is, for a given η, to iteratively assign the subchannels among the SUs to update $\{\mathcal{S}_k\}$, aiming at improving the minimum entry of $\{\mathcal{J}_k(\eta) \triangleq \widehat{J}_k^{(\mathcal{S}_k)}(\eta) \mid k \in \mathcal{K}\}$, till all the subchannels are assigned. Then approximate the value of $\underline{\widehat{J}}(\eta)$ with the resultant maximum entry of $\{\min_{k\in\mathcal{K}} \mathcal{J}_k(\eta)\}$ and adjust η accordingly. Repeat the above procedure till the stopping criterion is satisfied. Note that the this heuristic approach cannot ensure the convergence of η to any point including $\underline{\xi}_o$. A stopping criterion, such as stopping when $\eta^{(t+1)} \leq \eta^{(t)}$, is required in practice, where t is the iteration number. The detailed procedures of the proposed low-complexity algorithm are summarized in Table 5.2.

5.3 Average-EE-Based Spectrum Access

We address the average-EE-based spectrum access in this section. Since the sum of concave-convex ratios generally does not have any kind of concavity or monotonicity, there may be many local optima other than the global one, making the problem quite challenging.

5.3.1 Optimal Solution

Motivated by [86, 87, 88], we develop a concave envelope based B&B approach to optimally solve an equivalent form of the relaxed average-EE problem in (5.12). Moreover, we significantly simplify the approach by exploiting some underlying properties of the energy-efficient transmission.

To set up an equivalent form of problem (5.12), we first define some auxiliary variables: $\boldsymbol{\alpha} \triangleq [\alpha_1, \cdots, \alpha_K]^T, \boldsymbol{\beta} \triangleq [\beta_1, \cdots, \beta_K]^T \in \mathbb{R}^K$,

$$\mathcal{D} \triangleq \mathcal{B}_1^{(\mathcal{D})} \times \mathcal{B}_2^{(\mathcal{D})} \times \cdots \times \mathcal{B}_K^{(\mathcal{D})} = \left\{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \middle| (\alpha_k, \beta_k) \in \mathcal{B}_k^{(\mathcal{D})}\right\}, \tag{5.15}$$

$$\mathcal{B}_k^{(\mathcal{D})} \triangleq \left\{(\alpha_k, \beta_k) \mid \underline{\alpha}_k^{(\mathcal{D})} \leq \alpha_k \leq \overline{\alpha}_k^{(\mathcal{D})}, \underline{\beta}_k^{(\mathcal{D})} \leq \beta_k \leq \overline{\beta}_k^{(\mathcal{D})}\right\}. \tag{5.16}$$

And define some constant regions as follows:

$$\mathcal{B}_k^{(0)} \triangleq \left\{(\alpha_k, \beta_k) \mid 0 \leq \alpha_k \leq p_k^{\max}, 0 \leq \beta_k \leq \overline{r}_k^{\max}\right\}, \tag{5.17}$$

$$\mathcal{D}^{(0)} \triangleq \mathcal{B}_1^{(0)} \times \mathcal{B}_2^{(0)} \times \cdots \times \mathcal{B}_K^{(0)}, \tag{5.18}$$

where $\overline{r}_k^{\max} \triangleq \max\limits_{\mathbf{p}_k \in \mathcal{P}_{kn \in \mathcal{N}}} \sum qW \log_2(1 + p_{n,k}\gamma_{n,k})$ and $p_k^{\max} \triangleq \min(N\delta_k, p_{\max})$.

Then, we introduce an auxiliary problem that proves to be equivalent to (5.12) under a certain condition. For any given $\mathcal{D} \subseteq \mathcal{D}^{(0)}$ (i.e., $\mathcal{B}_k^{(\mathcal{D})} \subseteq \mathcal{B}_k^{(0)}$), we consider the following problem:

$$(\boldsymbol{\Phi}_o^{(\mathcal{D})}, \mathbf{P}_o^{(\mathcal{D})}, \boldsymbol{\alpha}_o^{(\mathcal{D})}, \boldsymbol{\beta}_o^{(\mathcal{D})}) \triangleq \arg\max_{\boldsymbol{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta}} \underbrace{\sum_{k \in \mathcal{K}} \frac{w_k \beta_k}{p_c + \rho \alpha_k}}_{\triangleq H_{\mathcal{D}}(\boldsymbol{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta})}, \tag{5.19a}$$

$$s.t. \quad \sum_{n \in \mathcal{N}} qW\varphi_{n,k} \log_2\left(1 + \frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}}\right) \geq \beta_k, \forall\, k \in \mathcal{K}, \tag{5.19b}$$

$$\sum_{n \in \mathcal{N}} p_{n,k} \leq \alpha_k, \forall\, k \in \mathcal{K}, \tag{5.19c}$$

$$\boldsymbol{\Phi} \in \Phi, \mathbf{P} \in \mathcal{P}, (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{D}, \tag{5.19d}$$

where $\boldsymbol{\Phi}_o^{(\mathcal{D})}$, $\mathbf{P}_o^{(\mathcal{D})}$, $\boldsymbol{\alpha}_o^{(\mathcal{D})}$, and $\boldsymbol{\beta}_o^{(\mathcal{D})}$ are the resultant optimal arguments. It is worth noting that even though the objective function, $H_{\mathcal{D}}(\boldsymbol{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta})$, is neither concave nor convex in α_k's

and β_k's, the feasible region of (5.19) proves to be either a non-empty convex set or an empty (convex) set.

The following theorem, proved in Appendix D.3, establishes an equivalent relationship between problem (5.19) and the relaxed average-EE problem in (5.12).

Theorem 5.5 *For $\mathcal{D} = \mathcal{D}^{(0)}$, a global optimal solution to problem (5.19) is also a global optimal solution to problem (5.12), and vice versa. That is,*

$$\overline{\xi}(\mathbf{\Phi}_o^{(\mathcal{D}^{(0)})}, \mathbf{P}_o^{(\mathcal{D}^{(0)})}) = \overline{\xi}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o),$$

and there exist $\boldsymbol{\alpha}', \boldsymbol{\beta}'$ satisfying $(\boldsymbol{\alpha}', \boldsymbol{\beta}') \in \mathcal{D}^{(0)}$ such that

$$H_{\mathcal{D}^0}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o, \boldsymbol{\alpha}', \boldsymbol{\beta}') = H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^{(0)})}, \mathbf{P}_o^{(\mathcal{D}^{(0)})}, \boldsymbol{\alpha}_o^{(\mathcal{D}^{(0)})}, \boldsymbol{\beta}_o^{(\mathcal{D}^{(0)})}).$$

Moreover, problems (5.19) and (5.12) have the same maximum value, i.e., $H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^{(0)})}, \mathbf{P}_o^{(\mathcal{D}^{(0)})}, \boldsymbol{\alpha}_o^{(\mathcal{D}^{(0)})}, \boldsymbol{\beta}_o^{(\mathcal{D}^{(0)})}) = \overline{\xi}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o)$.

Note that problems (5.12) and (5.19) do not necessarily have a unique global optimum, respectively. Hence, in general, we cannot simply claim $\mathbf{\Phi}_o^{(\mathcal{D}^{(0)})} = \overline{\mathbf{\Phi}}_o$ or $\mathbf{P}_o^{(\mathcal{D}^{(0)})} = \widetilde{\overline{\mathbf{P}}}_o$. From Theorem 5.5, we can solve (5.19) instead of directly dealing with (5.12) for the average-EE-based spectrum access.

Before presenting the B&B algorithm for solving (5.19), we first introduce some fundamentals for its two basic processes: bounding and branching. To compute upper bounds for the bounding process in the B&B algorithm, we approximate the K functions in the form of $h(\alpha, \beta) \triangleq \frac{\beta}{p_c + \rho\alpha}$ in the objective function of problem (5.19) by their concave envelopes, $h^{\mathrm{CE}}_{\mathcal{B}_k^{(\mathcal{D})}}(\alpha_k, \beta_k)$'s, which can be readily obtained from Appendix D.4. Since the concave envelope [89] of a function is its lowest (or tightest) concave overestimator, we can have a concave upper bound for $H_{\mathcal{D}}(\mathbf{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ in problem (5.19) as follows:

$$\widehat{H}_{\mathcal{D}}(\mathbf{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \triangleq \sum_{k \in \mathcal{K}} w_k h^{\mathrm{CE}}_{\mathcal{B}_k^{(\mathcal{D})}}(\alpha_k, \beta_k). \tag{5.20}$$

Then, the following problem is apparently a concave optimization problem:

$$(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})}) \triangleq \arg\max_{\mathbf{\Phi},\mathbf{P},\boldsymbol{\alpha},\boldsymbol{\beta}} \widehat{H}_{\mathcal{D}}(\mathbf{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \tag{5.21a}$$

$$s.t. \ \text{(5.19b)-(5.19d)}, \tag{5.21b}$$

which yields an upper bound on the maximum value of problem (5.19) and can be efficiently solved by convex optimization techniques, such as interior-point methods [76]. Then, for any $\mathcal{D} \subseteq \mathcal{D}^{(0)}$, we can obtain an upper bound and a feasible lower bound on the maximum of $H_{\mathcal{D}}(\mathbf{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ in (5.19) as follows:

$$\underbrace{H_{\mathcal{D}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})})}_{\text{a feasible lower bound}} \leq \underbrace{H_{\mathcal{D}}(\mathbf{\Phi}_o^{(\mathcal{D})}, \mathbf{P}_o^{(\mathcal{D})}, \boldsymbol{\alpha}_o^{(\mathcal{D})}, \boldsymbol{\beta}_o^{(\mathcal{D})})}_{\text{the maximum in } \mathcal{D}} \leq \underbrace{\widehat{H}_{\mathcal{D}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})})}_{\text{an upper bound}}, \tag{5.22}$$

if problem (5.19) is feasible for $\mathcal{D}$.

On the other hand, a proper branching/partition strategy is also required for the branching process in the B&B algorithm. For each survival branch/region (i.e., a region that needs further investigation), e.g., $\mathcal{D}$, the adopted branching strategy subdivides it into two branches/subregions, $\dot{\mathcal{D}}$ and $\ddot{\mathcal{D}}$, by bisecting the longest edge of $\boldsymbol{\alpha}$ of $\mathcal{D}$. The details of the branching strategy are put in Appendix D.5.

Based on (5.20)-(5.22) and the results in Appendices D.4 and D.5, we can develop a B&B algorithm to solve problem (5.19). The B&B algorithm for solving problem (5.19) with $\mathcal{D} = \mathcal{D}^{(0)}$ consists of three basic processes: bounding, pruning, and branching. For each survival branch/subregion (e.g., $\mathcal{D} \subseteq \mathcal{D}^{(0)}$), the bounding process computes an upper bound (e.g., $\widehat{H}_{\mathcal{D}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})})$) and a feasible lower bound (e.g., $H_{\mathcal{D}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})})$) on the maximum of (5.19) (e.g., $H_{\mathcal{D}}(\mathbf{\Phi}_o^{(\mathcal{D})}, \mathbf{P}_o^{(\mathcal{D})}, \boldsymbol{\alpha}_o^{(\mathcal{D})}, \boldsymbol{\beta}_o^{(\mathcal{D})})$) by solving (5.21) and plugging the resultant arguments (e.g., $\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})}$) into the original functions, respectively, if problem (5.21) is feasible for that branch. The pruning process prunes the branches that can already be excluded from further consideration. For a branch, where either problem (5.21) is infeasible or its upper bound is less than a feasible lower bound of any other branch, it can just be pruned at this point of search. The branching process iteratively subdivides

Table 5.3. Optimal Solution for the Average-EE-Based Spectrum Access based on the proposed B&B Algorithm.

1) (Bounding process) For each survival branch (e.g., $\mathcal{D}$), compute the upper bound (e.g., $\widehat{H}_{\mathcal{D}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})})$) by solving problem (5.21) and the lower bound (e.g., $H_{\mathcal{D}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})})$ by plugging the resultant arguments into the original functions, if problem (5.21) is feasible for the branch. Otherwise, notate the branch infeasible. Note that the initial survival branch is $\mathcal{D}^{(0)}$.

2) (Pruning process) Prune the infeasible branches and the branches whose upper bounds are less than the greatest lower bound of all branches (e.g., if $\widehat{H}_{\mathcal{D}'}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}')}, \widehat{\mathbf{P}}_o^{(\mathcal{D}')}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}')}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}')}) < \max_{\mathcal{D}} H_{\mathcal{D}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D})}, \widehat{\mathbf{P}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D})})$, then prune $\mathcal{D}'$).

3) (Branching process) Subdivide each survival branch (e.g., $\mathcal{D}$) into two branches (e.g., $\dot{\mathcal{D}}$ and $\ddot{\mathcal{D}}$) based on the branching strategy in Appendix D.5.

4) Repeat Steps 1)-3) till convergence or the stopping criterion is satisfied. Output the resultant subchannel and power allocation matrices, $\mathbf{\Phi}_o^{(\mathcal{D}^{(0)})}$ and $\mathbf{P}_o^{(\mathcal{D}^{(0)})}$, as a near-optimal solution (upper bound) for problem (5.7).

5) Round off the subchannel allocation matrix $\mathbf{\Phi}_o^{(\mathcal{D}^{(0)})}$ to a feasible one. Recalculate each user's transmit power vector to maximize their respective EE.

each survival rectangular region/branch (e.g., $\mathcal{D}$) into two subregions/branches (e.g., $\dot{\mathcal{D}}$ and $\ddot{\mathcal{D}}$) based on the branching strategy in Appendix D.5, creating a more refined partition of each survival branch that is still a possible candidate for the global optimum. The detailed procedures of the proposed B&B algorithm are summarized in Table 5.3.

The next theorem, proved in Appendix D.6, ensures the global convergence of the proposed B&B approach.

Theorem 5.6 *The proposed B&B algorithm converges to the optimal solution to problem (5.19) with $\mathcal{D} = \mathcal{D}^{(0)}$. More specifically, the difference of the maximum of all upper bounds and the maximum of the original problem, i.e.,*

$$\max_j \widehat{H}_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})}) - H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^{(0)})}, \mathbf{P}_o^{(\mathcal{D}^{(0)})}, \boldsymbol{\alpha}_o^{(\mathcal{D}^{(0)})}, \boldsymbol{\beta}_o^{(\mathcal{D}^{(0)})}),$$

is a decreasing function of i, where $\mathcal{D}^{(i,j)}$ denotes survival branch $j \in \{1, 2, \cdots\}$ after the i-th branching process. And the maximum of the upper bounds and lower bounds both

converge to the maximum of the original problem, i.e.,

$$\begin{aligned}
&\lim_{i\to\infty}\max_{j}\widehat{H}_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})},\widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})},\widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})},\widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})})\\
=&\lim_{i\to\infty}\max_{j}H_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})},\widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})},\widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})},\widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})})\\
=&H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^0)},\mathbf{P}_o^{(\mathcal{D}^0)},\boldsymbol{\alpha}_o^{(\mathcal{D}^0)},\boldsymbol{\beta}_o^{(\mathcal{D}^0)}).
\end{aligned}$$

In fact, it is unnecessary to set $\mathcal{D}$ as large as $\mathcal{D}^{(0)}$ to guarantee the equivalence of problems (5.19) and (5.12). We then further exploit the underlying properties of the energy-efficient transmission to significantly reduce the initial search space of $(\boldsymbol{\alpha},\boldsymbol{\beta})$ from $\mathcal{D}^{(0)}$ and thus to reduce the complexity of the proposed B&B algorithm. Next, we first give some definitions and then reveal some intrinsic properties of the energy-efficient transmission by Theorem 5.7 as proved in Appendix D.7.

For any given subchannel allocation indicator vector, $\boldsymbol{\varphi}_k \in \Phi_k$, let $\mathbf{p}_k(\boldsymbol{\varphi}_k) \triangleq [p_{1,k}(\boldsymbol{\varphi}_k),\cdots,p_{N,k}(\boldsymbol{\varphi}_k)]^T$ be the transmit power vector in $\mathcal{P}_k$ that maximizes the modified EE, $\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k)$. That is, $\mathbf{p}_k(\boldsymbol{\varphi}_k) \equiv \underset{\mathbf{p}_k\in\mathcal{P}_k}{\operatorname{argmax}}\,\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k)$. Define $p_k(\boldsymbol{\varphi}_k) \triangleq \sum_{n\in\mathcal{N}} p_{n,k}(\boldsymbol{\varphi}_k)$ and $\widetilde{r}_k(\boldsymbol{\varphi}_k) \triangleq \sum_{n\in\mathcal{N}} qW\varphi_{n,k}\log_2(1+\frac{p_{n,k}(\boldsymbol{\varphi}_k)\gamma_{n,k}}{\varphi_{n,k}})$.

Theorem 5.7 *For any two subchannel allocation indicator vectors, $\dot{\boldsymbol{\varphi}}_k \in \Phi_k$ and $\ddot{\boldsymbol{\varphi}}_k \in \Phi_k$, that satisfy the following component-wise inequality condition, i.e., $\dot{\boldsymbol{\varphi}}_k \preceq \ddot{\boldsymbol{\varphi}}_k$, we always have $p_k(\dot{\boldsymbol{\varphi}}_k) \le p_k(\ddot{\boldsymbol{\varphi}}_k) \le p_k(\mathbf{1}_N)$ and $\widetilde{r}_k(\dot{\boldsymbol{\varphi}}_k) \le \widetilde{r}_k(\ddot{\boldsymbol{\varphi}}_k) \le \widetilde{r}_k(\mathbf{1}_N)$, where $\preceq$ denotes the component-wise inequality and $\mathbf{1}_N \triangleq [\underbrace{1,1,\cdots,1}_{N}]^T$.*

Theorem 5.7 implies that, in energy-efficient transmission, a SU tends to transmit at a higher power and achieve a larger throughput when it occupies larger fractions on all subchannels. The next theorem, proved in Appendix D.8, helps shrink the original search space of $(\boldsymbol{\alpha},\boldsymbol{\beta})$ from $\mathcal{D}^{(0)}$ to a relatively small but sufficient extent for getting $(\boldsymbol{\alpha}_o^{(\mathcal{D}^0)},\boldsymbol{\beta}_o^{(\mathcal{D}^0)})$.

Theorem 5.8 *The optimal $(\boldsymbol{\alpha}_o^{(\mathcal{D}^0)},\boldsymbol{\beta}_o^{(\mathcal{D}^0)})$ to problem (5.19) with $\mathcal{D}=\mathcal{D}^{(0)}$ belongs to $\widetilde{\mathcal{D}}^{(0)}$.*

That is, $(\boldsymbol{\alpha}_o^{(\mathcal{J}^{(0)})}, \boldsymbol{\beta}_o^{(\mathcal{J}^{(0)})}) \in \widetilde{\mathcal{D}}^{(0)} \subseteq \mathcal{D}^{(0)}$*, where* $\widetilde{\mathcal{D}}^{(0)}$ *is defined as*

$$\widetilde{\mathcal{D}}^{(0)} \triangleq \widetilde{\mathcal{B}}_1^{(0)} \times \widetilde{\mathcal{B}}_2^{(0)} \times \cdots \times \widetilde{\mathcal{B}}_K^{(0)},$$
$$\widetilde{\mathcal{B}}_k^{(0)} \triangleq \left\{ (\alpha_k, \beta_k) \middle| 0 \leq \alpha_k \leq p_k(\mathbf{1}_N), 0 \leq \beta_k \leq \widetilde{r}_k(\mathbf{1}_N) \right\}.$$

Accordingly, $\widetilde{\mathcal{D}}^{(0)}$ is large enough to contain all the optimal solutions for problem (5.19) with $\mathcal{D} = \mathcal{D}^{(0)}$. Hence, Theorem 5.5 also holds for $\mathcal{D} = \widetilde{\mathcal{D}}^{(0)}$. Since the size of $\widetilde{\mathcal{D}}^{(0)}$ is about

$$\kappa \triangleq \prod_{k \in \mathcal{K}} \frac{p_k(\mathbf{1}_N)}{p_k^{\max}} \prod_{k \in \mathcal{K}} \frac{\widetilde{r}_k(\mathbf{1}_N)}{\widetilde{r}_k^{\max}}$$

of that of $\mathcal{D}^{(0)}$, we may at most reduce the complexity of the original B&B algorithm without losing optimality by a factor of κ if beginning searching with $\widetilde{\mathcal{D}}^{(0)}$ instead of $\mathcal{D}^{(0)}$. In some practical scenarios [79], κ can roughly be $(\frac{1}{8})^K$ with $\frac{p_k(\mathbf{1}_N)}{p_k^{\max}} \approx \frac{1}{4}$ and $\frac{\widetilde{r}_k(\mathbf{1}_N)}{\widetilde{r}_k^{\max}} \approx \frac{1}{2}$. Note that $\mathbf{p}_k(\mathbf{1}_N)$ proves to be unique from Theorem 5.3 and can be easily calculated following Theorem 5.4 and its subsequent approach. Thus, $p_k(\mathbf{1}_N)$ and $\widetilde{r}_k(\mathbf{1}_N)$ can also be easily obtained.

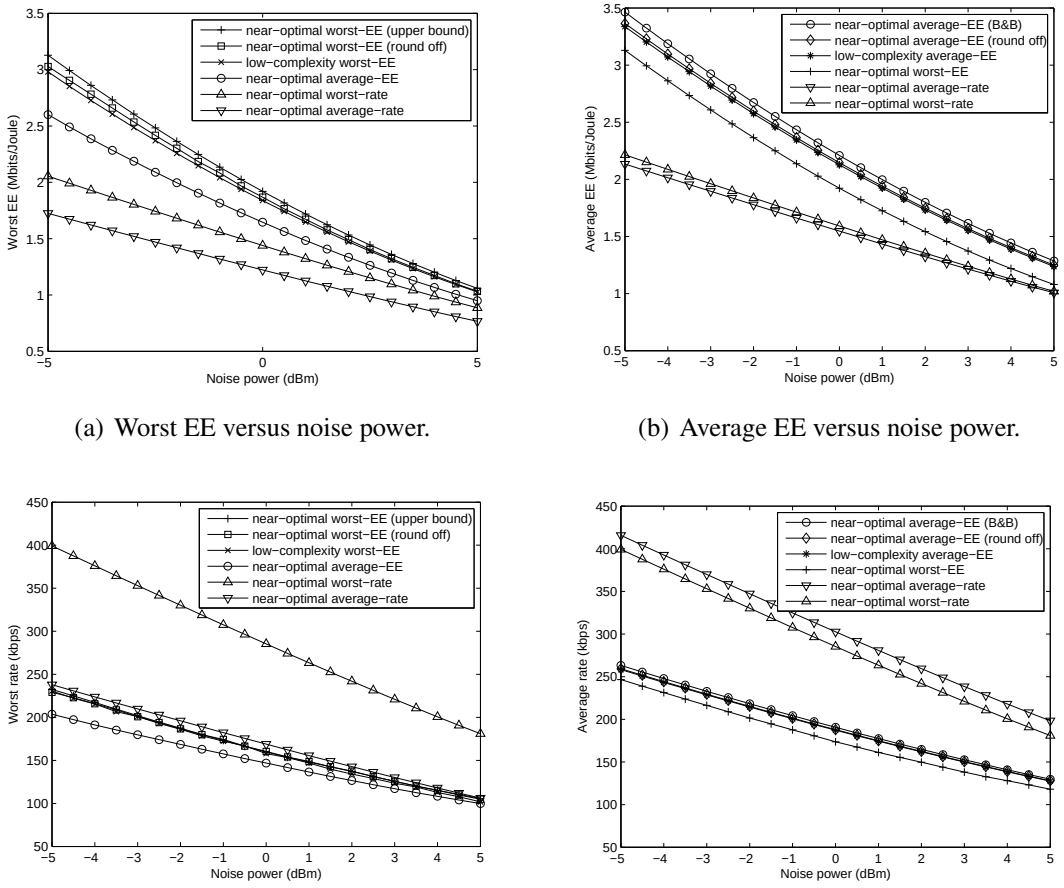
5.3.2 Low-Complexity Suboptimal Solution

To facilitate the development of an even low-complexity heuristic approach, we then present the following result for the average-EE-based case, which can be similarly proved as Theorem 5.3.

Theorem 5.9 *For any given subchannel allocation matrix,* $\boldsymbol{\Theta}_{\mathrm{v}}$*, the average EE,* $\overline{\mathcal{E}}^{(\boldsymbol{\Theta}_{\mathrm{v}})}(\mathbf{P}) \triangleq \sum_{k \in \mathcal{K}} \frac{w_k}{\sum_{k' \in \mathcal{K}} w_{k'}} \mathcal{E}_k^{(\boldsymbol{\theta}_k^{\mathrm{v}})}(\mathbf{p}_k)$*, is strictly quasiconcave in* $\mathbf{p}_k \in \mathcal{P}_k$ *and therefore has a unique global maximum. Moreover, the maximum of the average EE equals the weighted average of the maximum individual EE, i.e,* $\overline{\mathcal{E}}_o^{(\boldsymbol{\Theta}_{\mathrm{v}})} \triangleq \max_{\mathbf{P} \in \mathcal{P}} \overline{\mathcal{E}}^{(\boldsymbol{\Theta}_{\mathrm{v}})}(\mathbf{P}) = \sum_{k \in \mathcal{K}} \frac{w_k}{\sum_{k' \in \mathcal{K}} w_{k'}} \mathcal{E}_{k,o}^{(\boldsymbol{\theta}_k^{\mathrm{v}})}$.

The basic idea of the proposed low-complexity algorithm is to iteratively assign the subchannels among the SUs till all the subchannels are assigned. Each time only the SU with the maximum EE improvement if adding its best subchannel among the unassigned

subchannels acquires that subchannel. The procedures of the low-complexity heuristic algorithm for directly solving the original average-EE problem in (5.7) are summarized in Table 5.4.



(a) Worst EE versus noise power.

(b) Average EE versus noise power.

(c) Worst rate versus noise power.

(d) Average rate versus noise power.

Figure 5.1. Performance evaluation and comparison of the worst-EE-based, average-EE-based, worst-rate-based, and average-rate-based spectrum access strategies.

5.4 Numerical Results

In this section, we present simulation results to verify the benefits of the energy-efficient design and the performance of the developed algorithms. The frequency spacing of adjacent subcarriers is 15 kHz. For each SU, we assume that the maximum transmit power is 100 mW, the circuit power is 50 mW, and the drain efficiency of power amplifier is 38%. The channel realizations for all subchannles between the SU transmitters and PU receivers

Table 5.4. Low-Complexity Suboptimal Solution for the Average-EE-Based Spectrum Access.

1. Initialize the subchannel set and EE for each SU as an empty set and zero, i.e., $\mathcal{S}_k \leftarrow \varnothing$ and $\mathcal{E}_k \leftarrow 0, \forall k \in \mathcal{K}$, repsectively. Set the initial unassigned subchannel set as the set of all available subchannels, i.e., $\mathcal{N}_{us} \leftarrow \mathcal{N}$.

2. For each SU $k \in \mathcal{K}$, find its best subchannel among all the unassigned subchannels, i.e., $n_k^* \leftarrow \arg\max_{n\in\mathcal{N}_{us}} \gamma_{n,k}$, and calculate its maximum weighted EE under constraint (5.3) with subchannel n_k^* added, i.e., $w_k\mathcal{E}_{k,o}^{(\mathcal{S}_k\cup\{n_k^*\})}$. For the SU with the maximum weighted EE increament, i.e., SU $k^* \leftarrow \arg\max_{k\in\mathcal{K}}\left(w_k\mathcal{E}_{k,o}^{(\mathcal{S}_k\cup\{n_k^*\})} - \mathcal{E}_k\right)$, update its EE and assign it its best subchannel, i.e., $\mathcal{E}_{k^*} \leftarrow w_k\mathcal{E}_{k^*,o}^{(\mathcal{S}_{k^*}\cup\{n_{k^*}^*\})}$ and $\mathcal{S}_{k^*} \leftarrow \mathcal{S}_{k^*} \cup \{n_{k^*}^*\}$. Remove subchannel $n_{k^*}^*$ from the unassigned subchannel set, i.e. $\mathcal{N}_{us} \leftarrow \mathcal{N}_{us} \setminus \{n_{k^*}^*\}$.

3. Repeat Step 2 till all the subchannels are assigned, i.e., $\mathcal{N}_{us} = \varnothing$. Output the resultant subchannel sets and power allocation matrix, $\{\mathcal{S}_k\}$ and $\mathbf{P}$, as the suboptimal solution.

are assumed to be from independent Rayleigh fading with unit power, i.e., $\mathbb{E}(\bar{g}_k) = 1$. The interference power constraint, $\bar{I}_{th}$, is assumed to be 10 mW. The posteriori probability of mis-detection is $1 - q = 0.05$.

For comparison, we also consider spectral-efficient spectrum access strategies aiming at optimizing worst rate and average rate. The worst-rate-guaranteed and average-rate-based problems are formulated by replacing the power terms in the denominators of the worst-EE-guaranteed and average-EE-based problems with one, respectively. The near-optimal strategies are based on convex optimization, respectively. Note that in the figures, the true EEs and rates are used by normalizing the weighted EEs by their weights, although the optimizations are based on the weighted EEs and rates.

In the first scenario, we consider two SUs of equal weights ($w_1 = w_2 = 1$) accessing 10 available subchannels. The channel realizations for subchannels between the first SU's transmitter and receiver are from independent Rayleigh fading with unit power, while the other SU's channel realizations are from independent Rayleigh fading with power of two. From Figure 5.1(a), the worst EE of the proposed low-complexity worst-EE-guaranteed spectrum access approach is at least 95% of the upper bound on EE (achieved by Steps 1-3 in Table 5.1) and at least 98% of the round-off result (achieved by Steps 1-5 in Table 5.1).

From Figure 5.1(b), similar observations can be found for the proposed low-complexity suboptimal and the B&B-based near-optimal average-EE-based spectrum access. Moreover, from Figure 5.1(a) and 5.1(b), the continuous relaxation does provide a tight upper bound. From Figure 5.1(c) and 5.1(d), the energy-efficient spectrum access strategies do not always transmit at the maximum power.

In the second scenario, we consider four SUs with 40 available subchannels. The channel realizations for all subchannels between the first two SUs' transmitters and receivers are from independent Rayleigh fading with unit power, while the other two SUs' channel realizations are from independent Rayleigh fading with power of four. We investigate the impact of weights on the performance of worst-EE-based spectrum access by equal weights ($w_1 = w_2 = w_3 = w_4 = 1$) and unequal weights ($w_1 = w_2 = 1$ and $w_3 = w_4 = 0.4$) and compared their performance with that of the average-EE-based spectrum access with equal weights ($w_1 = w_2 = w_3 = w_4 = 1$). In Figure 5.2, we plot the average EE of SU 1 and SU 2, and the average EE of SU 3 and SU 4. It can be seen that, for the worst-EE-based in the case with equal weights, the SUs with lower average CNR have close EE compared to the SUs with higher CNR and impede the overall or average performance, which is a natural result of max - min optimization and absolute fairness. By giving the SUs with higher CNR smaller weights, they have much higher average EE compared to the SUs with lower CNR. Hence, by properly choosing the weights, the worst-EE-guaranteed spectrum access may also have a good balance between the overall or average performance and the worst performance.

Next, we evaluate the convergence performance of the low-complexity suboptimal worst-EE-based spectrum access. We also compare the performance of the proposed algorithm with a heuristic algorithm modified from the MUSA algorithm in [85]. The basic idea of the heuristic algorithm (and MUSA) is to iteratively assign the subchannels among the SUs, aiming at improving the minimum $\mathcal{E}_k$, till all the subchannels are assigned. Here we assume that all SUs have equal weights ($w_1 = w_2 = \cdots = w_K = 1$). The channels

between the SU transmitters and receivers are independent realizations of Rayleigh fading with unit power. Figure 5.3 plots the performance of the low-complexity suboptimal worst-EE-based spectrum access with different maximum iteration number, $iter_{\max} = 1, 2, 3, 4$, for doing Steps 1-4 in Table 5.2. It can be seen that the low-complexity approach converges quite fast for different number of SUs and subchannels, considering that on average at most four iterations lead to the best performance of low-complexity approach. Moreover, the low-complexity worst-EE-based spectrum access with $iter_{\max} = 2$ can trade some EE for data rate and thus provide a promising tradeoff between worst EE and worst rate. In addition, the heuristic algorithm [85] is slightly inferior to the proposed low-complexity algorithm in terms of EE and rate at a similar computational cost. However, compared with the proposed algorithms, the heuristic algorithm lacks the flexibility of trading off EE and rate by adjusting the iteration number.

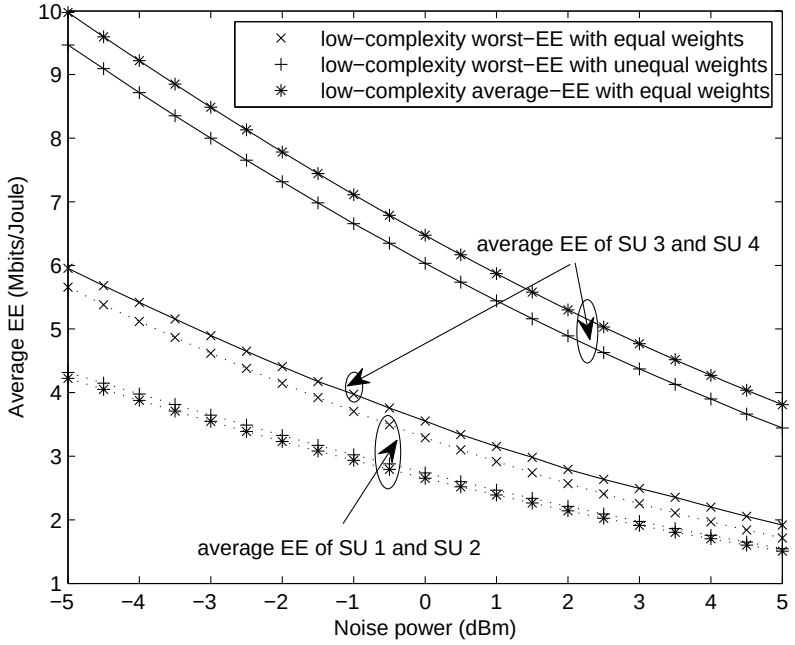


Figure 5.2. Average EE versus noise power.

Finally, we analyze the complexity of the proposed energy-efficient spectrum access strategies. Let $O(\mathcal{I})$ denote the complexity of getting the maximum EE for one SU with

a given set of assigned subchannels. To get the exact optimal performance for worst-EE-based and average-EE-based spectrum access, exhaustive search for all K^N combinations of subchannel assignment, each with a complexity of $K\mathcal{O}(\mathcal{I})$, is needed. Thus, the optimal complexity of the optimal approaches are $\mathcal{O}(K^{N+1}\mathcal{I})$. For every η, the low-complexity worst-EE-based approach takes $(K+N)\mathcal{O}(\mathcal{I})$ for solution. Then, the total complexity is $m_\eta(K+N)\mathcal{O}(\mathcal{I})$, where m_η denotes the number of iterations for different η. For the near-optimal worst-EE-based strategy, the complexity for every η is at least $\mathcal{O}(\frac{1}{\epsilon^2}KN)\mathcal{O}(\mathcal{I})$ for ϵ-optimality [67], which means the deviation from the true value is less than ϵ. Then, its total complexity is at least $m_\eta\mathcal{O}(\frac{1}{\epsilon^2}KN)\mathcal{O}(\mathcal{I})$. Similarly, the concave envelope based B&B approach needs $\mathcal{O}(\frac{1}{\epsilon^2}KN\mathcal{I}m_{\mathcal{D}^{(0)}})$ till convergence, where $m_{\widetilde{\mathcal{D}}^{(0)}}$ denotes the number of visited branches starting with $\widetilde{\mathcal{D}}^{(0)}$. The low-complexity average-EE-based approach takes a complexity of $K\mathcal{O}(\mathcal{I})$ to assign each subchannel, resulting in a total complexity of $KN\mathcal{O}(\mathcal{I})$. It is obvious that the suboptimal approaches have much lower complexity than that of optimal and near-optimal approaches. The complexity is summarized in Table 5.5.

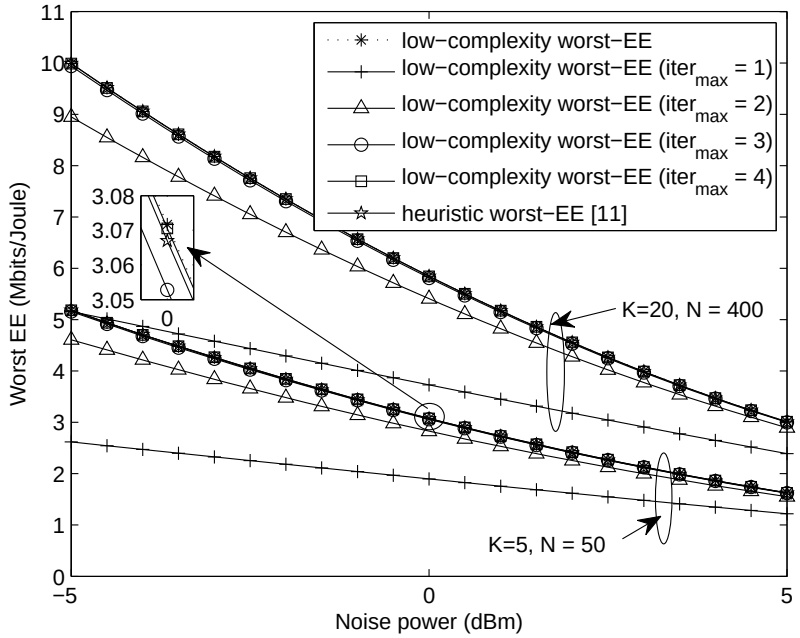
Table 5.5. Complexity Comparison for Energy-Efficient Spectrum Access Strategies

Energy-Efficient Spectrum Access	Complexity
worst-EE-based: *optimal* by exhaustive search	$\mathcal{O}(K^{N+1}\mathcal{I})$
worst-EE-based: *near-optimal* (in Table 5.1)	$\mathcal{O}(\frac{1}{\epsilon^2}KN\mathcal{I}m_\eta)$
worst-EE-based: *suboptimal* (in Table 5.2)	$\mathcal{O}((K+N)\mathcal{I}m_\eta)$
average-EE-based: *optimal* by exhaustive search	$\mathcal{O}(K^{N+1}\mathcal{I})$
average-EE-based: *near-optimal* (concave envelope based B&B approach)	$\mathcal{O}(\frac{1}{\epsilon^2}KN\mathcal{I}m_{\widetilde{\mathcal{D}}^{(0)}})$
average-EE-based: *suboptimal* (in Table 5.4)	$\mathcal{O}(KN\mathcal{I})$

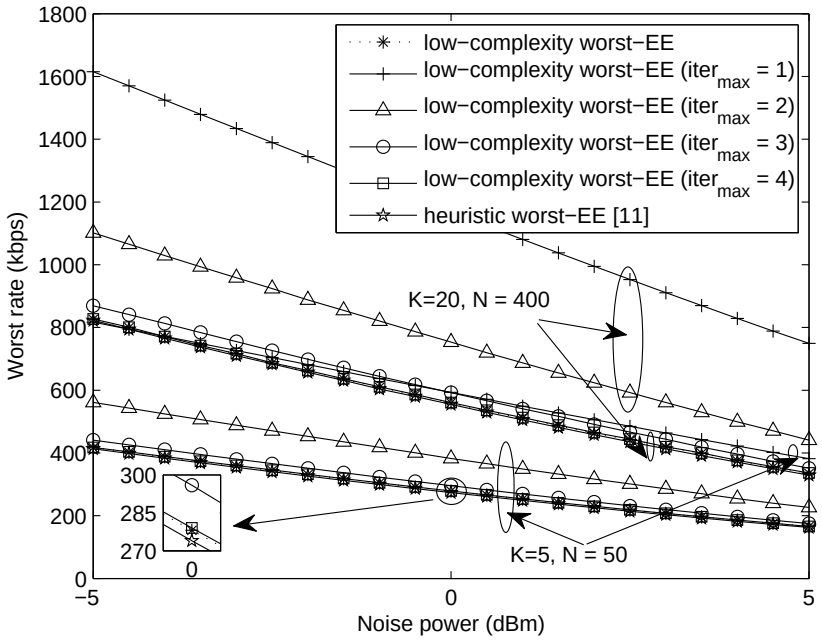
5.5 Conclusion

In this chapter, we have studied two types of energy-efficient spectrum access problems in an OFDM-based CR network. Both problems belong to the nonconvex integer combinatorial fractional program and are NP-hard. The worst-EE-based problem proves to be quasiconcave after continuous and concave relaxation and can thus be near-optimally solved.

We also develop a suboptimal alternative to further reduce complexity. The average-EE-based problem may have many local optima even after continuous and concave relaxation due to its sum-of-ratios structure. We first convert it into an equivalent form and introduce a concave envelope based B&B approach to find its optimal solution. We then exploit some underlying properties of the energy-efficient transmission to speed up the convergence of the B&B approach. We also develop a low-complexity suboptimal approach. Simulation results show that the energy-efficient spectrum access strategies significantly improve EE compared to the conventional spectral-efficient spectrum access while the low-complexity suboptimal approaches can well balance the performance and complexity.



(a) Worst EE versus noise power.



(b) Worst rate versus average power.

Figure 5.3. Convergence performance evaluation of the low-complexity suboptimal worst-EE-based spectrum access.

CHAPTER 6

ENERGY-EFFICIENT OFDMA-BASED TWO-WAY RELAY

In this chapter, we concentrate on energy-efficient joint power and subchannel allocation, including active subchannel selection, for OFDMA-based two-way relay networks, while providing proportional fairness in EE among different terminal pairs. Even though our analysis and methods are based on AF relay, they can be extended to DF or *compress-and-forward* (CF) relay. The joint power and subchannel allocation problem is mixed-integer combinatorial and further turns to be nonconvex and NP-hard. We first find an upper bound using continuous relaxation and the Lagrange dual method. To reduce the computational complexity, we then develop a suboptimal alternative by exploiting the the hidden concavity and the pseudoconcavity in the subproblems for any fixed subchannel assignment and suggesting an EE-oriented sequential subchannel assignment policy. Moreover, we discover the sufficient condition for early termination of the sequential subchannel assignment without losing the EE optimality. Simulation results demonstrate the energy-efficient OFDMA-based two-way relay can achieve much larger EE utility while provisioning proportional fairness in EE among different terminal pairs compared to its spectral-efficient counterpart. Furthermore, the proposed suboptimal approach can achieve a satisfying EE performance at a relatively low complexity.

The rest of the chapter is organized as follows. In Section 6.1, we describe the system model, including the two-way relay mechanism and the power consumption model, and formulate the energy-efficient resource allocation problem for the OFDMA-based two-way relay. In Section 6.2, an upper bound for the optimization problem is obtained. In Section 6.3, a low-complexity suboptimal scheme is developed. Then, simulation results are presented in Section 6.4. Finally, we conclude this chapter in Section 6.5.

6.1 System Model and Problem Formulation

In this section, we present the system model and formulate the energy-efficient resource allocation problem for the OFDMA-based two-way relay network.

6.1.1 System Model

We consider a centralized wireless OFDMA-based two-way relay network shown in Figure 6.1, which consists of one relay and N pair of terminals, all half-duplex. Terminal pair n is formed by terminals $T_n^{(j)}$'s, $j \in \mathcal{J} \triangleq \{1,2\}$. All the terminal pairs denoted by $\mathcal{N} \triangleq \{1,\cdots,n,\cdots,N\}$ orthogonally share I subchannels denoted by $\mathcal{I} \triangleq \{1,\cdots,i,\cdots,I\}$ and exchange information within each pair via the intermediate relay. Let $\mathcal{H}_n$ denote the subchannel index set for terminal pair n. Then, $\mathcal{H}_n \bigcap \mathcal{H}_{n'} = \varnothing$, for $n \neq n' \in \mathcal{N}$ and $\bigcup_{n\in\mathcal{N}} \mathcal{H}_n \subseteq \mathcal{I}$.

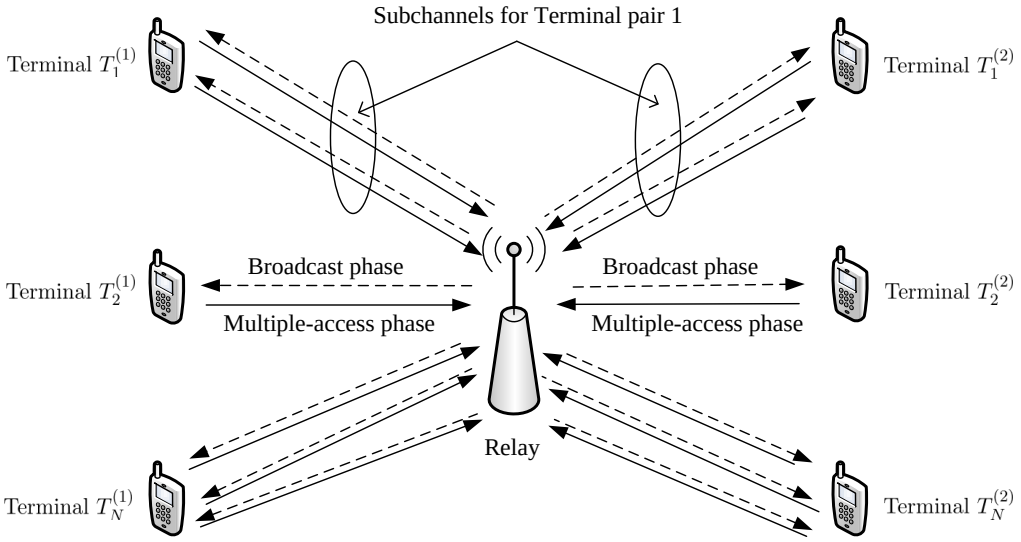


Figure 6.1. Structure of OFDMA-based two-way relay.

6.1.1.1 Two-Way AF Relay

We study *time-division duplexing* (TDD) two-phase AF relay strategy, which consists of a multiple-access phase and a subsequent broadcast phase of the same duration. Channels for two consecutive phases are assumed to be reciprocal. Denote $h_{n,i}^{(j)}$ to be the channel coefficient of the link between terminal $T_n^{(j)}$ and the relay on subchannel i.

In the multiple-access phase, all terminals simultaneously transmit signals intended for their pairing partners to the relay over the assigned subchannels. If subchannel i is assigned to terminal pair n or say $i \in \mathcal{H}_n$, the received signal at the relay on that subchannel is

$$y_i = \sum_{j \in \mathcal{J}} \sqrt{p_{n,i}^{(j)}} h_{n,i}^{(j)} x_{n,i}^{(j)} + z_i,$$

where $x_{n,i}^{(j)}$ and $p_{n,i}^{(j)}$ are the transmitted symbol and transmit power of terminal $T_n^{(j)}$ on subchannel i, respectively. $z_i \sim \mathcal{CN}(0,1)$ is the additive white Gaussian noise on subchannel i at the relay.

After receiving the signal, y_i, the relay, which transmits at a power of q_i on subchannel i, amplifies it by a scalar of $\frac{\sqrt{q_i}}{1+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2}$, and then broadcasts the amplified signal to terminal pair n over the same subchannel i in the broadcast phase. The received signals at terminal $T_n^{(j)}$ on subchannel i in the broadcast phase is

$$y_{n,i}^{(j)} = h_{n,i}^{(j)} \frac{\sqrt{q_i}}{1 + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2} y_i + z_{n,i}^{(j)},$$

where $z_{n,i}^{(j)} \sim \mathcal{CN}(0,1)$ is the additive white Gaussian noise at terminal $T_n^{(j)}$ on subchannel i. The received *signal-to-noise ratio* (SNR) at terminal $T_n^{(j)}$ on subchannel i after canceling the self-interference that comes from $x_{n,i}^{(j)}$ is

$$\gamma_{n,i}^{(j)} = \frac{p_{n,i}^{(3-j)} q_i |h_{n,i}^{(1)}|^2 |h_{n,i}^{(2)}|^2}{1 + q_i|h_{n,i}^{(j)}|^2 + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2}.$$

The achievable rate region for terminal pair n on subchannel i for the AF relay is [15, 55, 16]

$$\mathcal{R}_{n,i}\left(\mathbf{p}_{n,i}\right) = \left\{ \mathbf{r}_{n,i} \in \mathbb{R}_+^2 \;\middle|\; r_{n,i}^{(j)} \le \frac{1}{2} \log_2\left(1 + \gamma_{n,i}^{(j)}\right) \right\}, \tag{6.1}$$

where $\mathbf{p}_{n,i} \triangleq \left(p_{n,i}^{(1)}, p_{n,i}^{(2)}\right)^T$ and $\mathbf{r}_{n,i} \triangleq \left(r_{n,i}^{(1)}, r_{n,i}^{(2)}\right)^T$. $r_{n,i}^{(j)}$ represents the received data rate for terminal $T_n^{(j)}$ on subchannel i. The factor of $\frac{1}{2}$ is because of the use of two orthogonal time-slots for relay.

The overall received data rate for terminal $T_n^{(j)}$ and the overall data rate for terminal pair n can be expressed as

$$r_n^{(j)} = \sum_{i \in \mathcal{I}} \varphi_{n,i} r_{n,i}^{(j)} \;\text{ and }\; r_n = r_n^{(1)} + r_n^{(2)},$$

respectively, where $\varphi_{n,i} \in \{1, 0\}$ is a binary subchannel assignment indicator informing whether subchannel i is assigned to terminal pair n (by $\varphi_{n,i} = 1$) or not (by $\varphi_{n,i} = 0$).

6.1.1.2 Power Consumption Model

The total power consumption of terminal $T_n^{(j)}$ is[1]

$$p_n^{(j),tot} = \frac{1}{2}\rho \sum_{i\in\mathcal{I}} p_{n,i}^{(j)} + p_c^S + p_c^D \sum_{i\in\mathcal{I}} \varphi_{n,i},$$

where $\frac{1}{\rho} > 1$ denotes the amplifier efficiency. $p_c^S > 0$ represents the static circuit power that is independent of the number of active subchannels at a terminal while $p_c^D \geq 0$ represents the dynamic/extra circuit power per active subchannel [90, 57]. Note that if p_c^D is set zero, the power consumption model above is reduced to the one in [91, **?**, 92]. Then, the total power consumption of terminal pair n is

$$p_n^{tot} = \sum_{j\in\mathcal{J}} p_n^{(j),tot} = 2p_c^D \sum_{i\in\mathcal{I}} \varphi_{n,i} + 2p_c^S + \frac{1}{2}\rho \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} p_{n,i}^{(j)}.$$

6.1.2 Problem Formulation

In this chapter, the EE of a terminal pair characterizes the number of exchanged information bits per Joule of energy consumed at that terminal pair. Then, the EE of terminal pair n is

$$\mathcal{E}_n \triangleq \frac{r_n}{p_n^{tot}} = \frac{\sum\limits_{i\in\mathcal{I}} \varphi_{n,i}(r_{n,i}^{(1)} + r_{n,i}^{(2)})}{2p_c^D \sum\limits_{i\in\mathcal{I}} \varphi_{n,i} + 2p_c^S + \frac{1}{2}\rho \sum\limits_{i\in\mathcal{I}} (p_{n,i}^{(1)} + p_{n,i}^{(2)})}.$$

We aim to optimize the aggregated EE utility of all terminal pairs with proportional fairness in EE of different terminal pairs. A typical application scenario is where the user terminals without external energy sources want to exchange information as much as possible via a relay station with external energy source. Apparently, the user terminal pairs prefer high EE to exchange more information while the relay station does not have to. For simplicity, we assume fixed transmit powers on the active subchannels at the relay, i.e.,

[1] Note that the factor of $\frac{1}{2}$ is because each terminal only transmits at the multiple-access phase.

$\mathbf{q} \triangleq (q_i)_{i=1}^{I}$, is predetermined, and focus on the subchannel and power allocation, including active subchannel selection, for the terminals. Mathematically, the energy-efficient resource allocation provisioning proportional fairness in EE for the OFDMA-based two-way AF relay can be formulated as follows:

$$\max_{\boldsymbol{\varphi}, (\overline{\mathbf{p}}_n)_{n=1}^{N}} \sum_{n\in\mathcal{N}} \ln \mathcal{E}_n, \tag{6.2a}$$

$$s.t. \quad \mathbf{r}_{n,i} \in \mathcal{R}_{n,i}\Big(\mathbf{p}_{n,i}\Big), \forall n \in \mathcal{N}, \tag{6.2b}$$

$$\mathbf{1}_{2I}^{T} \cdot \overline{\mathbf{p}}_n \leq p_n^{\max},\ \overline{\mathbf{p}}_n \succeq \mathbf{0}_{2I}, \forall n \in \mathcal{N}, \tag{6.2c}$$

$$\sum_{n\in\mathcal{N}} \varphi_{n,i} \leq 1,\ \varphi_{n,i} \in \{0,1\}, \forall i \in \mathcal{I}, n \in \mathcal{N}, \tag{6.2d}$$

where $\ln(\cdot)$ denotes the natural logarithm function. $\overline{\mathbf{p}}_n \triangleq (\mathbf{p}_{n,i})_{i=1}^{I}$ is the collection of all $\mathbf{p}_{n,i}$'s for terminal pair n and $p_n^{\max}$ is the maximum total transmit power of terminal pair n.[2] $\boldsymbol{\varphi} \triangleq (\varphi_{n,i})_{u=1,i=1}^{N,I}$ is the matrix of subchannel assignment indicators. Constraints in (6.2b) are due to a simple fact that every feasible rate vector of one terminal pair on a subchannel lies in the rate region. Constraints in (6.2c) impose the nonnegative and maximum transmit power constraints for all the terminal pairs while constraints in (6.2d) guarantees that each subchannel is assigned to at most one terminal pair. Note that $\sum_{n\in\mathcal{N}} \varphi_{n,i} = 0$ implies subchannel i is inactive.

Remark 1: If DF or CF relay is adopted instead of AF, similar energy-efficient OFDMA-based two-way relay problems can be formulated by replacing the rate region for AF in (6.2b) by the rate regions for DF [15, 55, 16] and CF [16], respectively.

6.2 Performance Bound for Energy-Efficient OFDMA-Based Two-Way AF Relay

In this section, we investigate the optimal energy-efficient resource allocation for the OFDMA-based two-way AF relay.

[2]Similar to the literature limiting the total transmit power of the terminal pair and the relay [54, 93] or minimizing the total transmit power of them [17], we restrict the sum of the transmit powers of each terminal pair. In general, the resultant EE is an upper bound on that of the corresponding problem with individual transmit power constraints for each terminal pair.

For the energy-efficient resource allocation problem in (6.2), constraints in (6.2b) can be safely removed by letting $r_{n,i}^{(j)} = \frac{1}{2}\log_2\left(1+\gamma_{n,i}^{(j)}\right)$, which is on the outer boundary of the rate region, and expressing the EE of terminal pair k, $\mathcal{E}_k$, as

$$\widehat{\mathcal{E}_n} = \frac{\sum\limits_{i\in\mathcal{I}}\varphi_{n,i}\left(\log_2\left(1+\gamma_{n,i}^{(1)}\right)+\log_2\left(1+\gamma_{n,i}^{(2)}\right)\right)}{4p_c^D\sum\limits_{i\in\mathcal{I}}\varphi_{n,i}+4p_c^S+\rho\sum\limits_{i\in\mathcal{I}}(p_{n,i}^{(1)}+p_{n,i}^{(2)})}. \tag{6.3}$$

Then, problem (6.2) can be equivalently recast as

$$\max_{\boldsymbol{\varphi},(\overline{\mathbf{p}}_n)_{n=1}^N} \sum_{n\in\mathcal{N}} \ln\widehat{\mathcal{E}_n}, \tag{6.4a}$$

$$s.t.\quad \mathbf{1}_{2I}^T\cdot\overline{\mathbf{p}}_n \le p_n^{\max},\ \overline{\mathbf{p}}_n \ge \mathbf{0}_{2I}, \forall n\in\mathcal{N}, \tag{6.4b}$$

$$\sum_{n\in\mathcal{N}}\varphi_{n,i}\le 1,\ \varphi_{n,i}\in\{0,1\}, \forall i\in\mathcal{I}, n\in\mathcal{N}. \tag{6.4c}$$

Next, we briefly discuss the general mathematical properties of problems (6.2) and (6.4). First of all, the rate region in (6.1) is nonconvex/nonconcave in $\mathbf{p}_{n,i}$ because $\log_2\left(1+\gamma_{n,i}^{(j)}\right)$ is nonconvex/nonconcave in $\mathbf{p}_{n,i}$. Hence, the entire constraint set of (6.2b)-(6.2d) itself is nonconvex/nonconcave for problem (6.2) while the constraint set of (6.4b)-(6.4c) as well as the objective function $\sum_{n\in\mathcal{N}}\ln\widehat{\mathcal{E}_n}$ are both nonconvex/nonconcave for problem (6.4). Hence, problems (6.2) and (6.4) are generally nonconvex/nonconcave and are NP-hard for optimal solutions. To approach the optimal solution to problem (6.4), we can utilize Lagrange dual method [59, 58, **?**, 76] to solve its dual problem, based on the following equivalent form of $\sum_{n\in\mathcal{N}}\ln\widehat{\mathcal{E}_n}$:

$$\sum_{n\in\mathcal{N}}\ln\widehat{\mathcal{E}_n} = \sum_{n\in\mathcal{N}}\ln\left(\sum_{i\in\mathcal{I}}\varphi_{n,i}\left(\log_2\left(1+\gamma_{n,i}^{(1)}\right)+\log_2\left(1+\gamma_{n,i}^{(2)}\right)\right)\right) - \sum_{n\in\mathcal{N}}\ln\left(4p_c^D\sum_{i\in\mathcal{I}}\varphi_{n,i}+4p_c^S+\rho\sum_{i\in\mathcal{I}}(p_{n,i}^{(1)}+p_{n,i}^{(2)})\right),$$

and the continuous relaxation of $\varphi_{n,i}\in\{0,1\}$ into $\varphi_{n,i}\in[0,1]$ as in [**?**]. We omit the details of the Lagrange dual method due to its somewhat standard procedure. In general, the duality gap between the dual optimal and the primal optimal is nonzero for this problem. Thus, the dual problem provides an upper bound on the optimal solution for problem (6.4).

Remark 2: Unlike the AF relay case, the rate regions for the DF and CF relay both prove to be concave in $\mathbf{p}_{n,i}$ [55, 16]. Unfortunately, even though each EE term, $\mathcal{E}_n$, is in the form

of concave-convex ratio [94], the sum of the coupled concave-convex ratios is generally nonconvex/noncave. Therefore, the DF or CF problem is also nonconvex/nonconcave like the AF case.

6.3 Low-Complexity Energy-Efficient OFDMA-Based Two-Way AF Relay

The complexity of dual method can be high and thus computationally infeasible in practice. To facilitate the application of energy-efficient OFDMA-based two-way AF relay, we focus on low-complexity suboptimal solutions for problem (6.4) in this section.

6.3.1 Optimal Energy-Efficient Power Allocation for Fixed Subchannel Allocation

To get insights for developing a low-complexity approach, we first study the optimal energy-efficient power allocation for fixed subchannel allocation, i.e., $\{\mathcal{H}_n | n \in \mathcal{N}\}$ is predetermined.

For any given $\{\mathcal{H}_n | n \in \mathcal{N}\}$ that allows problem (6.4) to have at least one feasible solution, problem (6.4) can be naturally decoupled into N power alligation subproblems, each corresponds to one terminal pair, as follows:

$$\textit{Subproblem } n: \quad \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)} \triangleq \max_{\overline{\mathbf{p}}_n} \widehat{\mathcal{E}}_n, \tag{6.5a}$$

$$s.t. \quad (6.4b), \tag{6.5b}$$

where $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}$ represents the optimal EE. Note that $\varphi_{n,i} = 1$ if $i \in \mathcal{H}_n$; otherwise, $\varphi_{n,i} = 0$.

To solve problem (6.5), we first reveal some properties of the problem by the following theorem proved in Appendix E.1.

Theorem 6.1 *For any $\xi_n \geq 0$,*

$$T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n) \triangleq \sum_{i\in\mathcal{H}_n}\left(\log_2\left(1+\gamma_{n,i}^{(1)}\right)+\log_2\left(1+\gamma_{n,i}^{(2)}\right)\right) - \xi_n\left(4p_c^D|\mathcal{H}_n| + 4p_c^S + \rho\sum_{i\in\mathcal{H}_n}(p_{n,i}^{(1)} + p_{n,i}^{(2)})\right),$$

under constraints in (6.5b) satisfies

$$T_n^{(\mathcal{H}_n)}(\xi_n) \triangleq \max_{\overline{\mathbf{p}}_n} T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n) = 0 \quad \text{iff} \quad \xi_n = \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}.$$

Moreover, $T_n^{(\mathcal{H}_n)}(\xi_n)$ strictly decreases with $\xi_n \geq 0$.

Theorem 6.1 suggests a framework for solving problem (6.5). According to Theorem 6.1, if $T_n^{(\mathcal{H}_n)}(\xi_n) > 0$, then $\xi_n < \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}$; if $T_n^{(\mathcal{H}_n)}(\xi_n) < 0$, then $\xi_k > \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}$. Therefore, provided that the value of $T_n^{(\mathcal{H}_n)}(\xi_n)$ is available for any $\xi_n \geq 0$, $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}$ can be easily attained by various root-finding methods, such as the Newton-Raphson method [95], to get the root of $T_n^{(\mathcal{H}_n)}(\xi_n) = 0$. However, $T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n)$ is generally not concave or convex in $\overline{\mathbf{p}}_n$ [17], making it difficult to maximize $T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n)$ directly. To obtain $T_n^{(\mathcal{H}_n)}(\xi_n)$ for any given $\xi_n \geq 0$, we focus on the following problem:

$$\mathcal{T}_n^{(\mathcal{H}_n)}(\xi_n) \triangleq \max_{\mathbf{p}_n} \sum_{i\in\mathcal{H}_n} \Big(C_{n,i}(p_{n,i}) - \xi_n \rho p_{n,i}\Big), \tag{6.6a}$$

$$s.t.\ \ \mathbf{1}_{|\mathcal{H}_n|}^T \cdot \mathbf{p}_n \leq p_n^{\max},\ \mathbf{p}_n \geq \mathbf{0}_{|\mathcal{H}_n|}, \tag{6.6b}$$

where $\mathbf{p}_n \triangleq (p_{n,i})_{i\in\mathcal{H}_n}$ and $p_{n,i} \triangleq p_{n,i}^{(1)} + p_{n,i}^{(2)}$. Let $\mathbf{p}_n^* \triangleq (p_{n,i}^*)_{i\in\mathcal{H}_n}$ denote one optimal $\mathbf{p}_n$ for problem (6.6). $C_{n,i}(p_{n,i})$ gives the maximum sum rate for terminal pair n on subchannel i for a given transmit power, $p_{n,i}$, on that subchannel, and can be expressed as follows:

$$C_{n,i}(p_{n,i}) \triangleq \max_{\substack{\mathbf{1}_2^T\cdot\mathbf{p}_{n,i}=p_{n,i},\\ \mathbf{p}_{n,i}\geq\mathbf{0}_2}} \Big(\log_2\Big(1+\gamma_{n,i}^{(1)}\Big)+\log_2\Big(1+\gamma_{n,i}^{(2)}\Big)\Big). \tag{6.7}$$

Let $\mathbf{p}_{n,i}^*(p_{n,i}) \triangleq (p_{n,i}^{(1),*}(p_{n,i}), p_{n,i}^{(2),*}(p_{n,i}))^T$ denote one optimal $\mathbf{p}_{n,i}$ for problem (6.7). Apparently, we can alternatively solve $\mathcal{T}_n^{(\mathcal{H}_n)}(\xi_n)$ to get $T_n^{(\mathcal{H}_n)}(\xi_n)$ since $T_n^{(\mathcal{H}_n)}(\xi_n) \equiv \mathcal{T}_n^{(\mathcal{H}_n)}(\xi_n) - 4\xi_n(p_c^D|\mathcal{H}_n| + p_c^S)$.

To solve the problems in (6.6)-(6.7), we introduce some properties of $C_{n,i}(p_{n,i})$ and $\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i}) \triangleq C_{n,i}(p_{n,i}) - \xi_n\rho p_{n,i}$, which are summarized in Theorem 6.2 proved in Appendix E.2 and Theorem 6.3, respectively.

Theorem 6.2 *For any given $q_i > 0$, $C_{n,i}(p_{n,i})$ has the following properties:*

1. *$C_{n,i}(p_{n,i})$ is strictly concave and strictly increasing in $p_{n,i} \geq 0$.*

2. *The following conditions are necessary for* $\mathbf{p}_{n,i}^*(p_{n,i})$*:*

$$\begin{cases} f_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i})) \le 0 & \text{if} \quad \mathbf{p}_{n,i}^*(p_{n,i}) = (p_{n,i}, 0)^T \\ f_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i})) \ge 0 & \text{if} \quad \mathbf{p}_{n,i}^*(p_{n,i}) = (0, p_{n,i})^T \\ f_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i})) = 0 & \text{if} \quad \mathbf{p}_{n,i}^*(p_{n,i}) \succ \mathbf{0}_2, \end{cases} \tag{6.8}$$

where $f_{n,i}(\mathbf{x})$ *is defined in (6.9).*

3. *The derivative of* $C_{n,i}(p_{n,i})$ *w.r.t* $p_{n,i}$ *satisfies*

$$d_{n,i}(p_{n,i}) \triangleq \frac{\mathrm{d}C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}} = \max\left\{ \left.\frac{\partial R_{n,i}(\mathbf{p}_{n,i})}{\partial p_{n,i}^{(1)}}\right|_{\mathbf{p}_{n,i}=\mathbf{p}_{n,i}^*(p_{n,i})}, \left.\frac{\partial R_{n,i}(\mathbf{p}_{n,i})}{\partial p_{n,i}^{(2)}}\right|_{\mathbf{p}_{n,i}=\mathbf{p}_{n,i}^*(p_{n,i})} \right\},$$

where $R_{n,i}(\mathbf{p}_{n,i}) \triangleq \log_2\left(1+\gamma_{n,i}^{(1)}\right)+\log_2\left(1+\gamma_{n,i}^{(2)}\right)$.

$$f_{n,i}(\mathbf{x}) = \frac{1+q_i|h_{n,i}^{(2)}|^2+x_1(q_i|h_{n,i}^{(1)}|^2|h_{n,i}^{(2)}|^2+|h_{n,i}^{(1)}|^2)+x_2|h_{n,i}^{(2)}|^2}{1+q_i|h_{n,i}^{(1)}|^2+x_2(q_i|h_{n,i}^{(1)}|^2|h_{n,i}^{(2)}|^2+|h_{n,i}^{(2)}|^2)+x_1|h_{n,i}^{(1)}|^2} \cdot \frac{1+q_i|h_{n,i}^{(2)}|^2+x_1|h_{n,i}^{(1)}|^2+x_2|h_{n,i}^{(2)}|^2}{1+q_i|h_{n,i}^{(1)}|^2+x_1|h_{n,i}^{(1)}|^2+x_2|h_{n,i}^{(2)}|^2} - \frac{1+(q_i+p_{n,i})|h_{n,i}^{(2)}|^2}{1+(q_i+p_{n,i})|h_{n,i}^{(1)}|^2}. \tag{6.9}$$

Aside from revealing the concavity of $C_{n,i}(p_{n,i})$, Theorem 6.2 also implicitly offers an efficient way to solve problem (6.7) or say to get $\mathbf{p}_{n,i}^*(p_{n,i})$. Since $f_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i})) = 0$ is a quadratic equation of $p_{n,i}^{(1),*}(p_{n,i})$ and $p_{n,i}^{(2),*}(p_{n,i})$ and $p_{n,i}^{(1),*}(p_{n,i}) + p_{n,i}^{(2),*}(p_{n,i}) = p_{n,i}$, there are at most two valid solutions to $\mathbf{p}_{n,i}^*(p_{n,i})$ for the third case in (6.8). Counting the other two cases in (6.8), there are at most four possibilities of $\mathbf{p}_{n,i}^*(p_{n,i})$ in total. Therefore, $\mathbf{p}_{n,i}^*(p_{n,i})$ can be readily determined by ruling out the invalid or inferior ones. Accordingly, $d_{n,i}(p_{n,i})$ can also be easily calculated. Moreover, based on Theorem 6.2, the following theorem is immediate.

Theorem 6.3 *For any given* $\xi_n \ge 0$ *and* $q_n > 0$, $\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})$ *has the following properties:*

1. $\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})$ *is strictly concave in* $p_{k,n} \ge 0$.

2. *The derivative of* $\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})$ *w.r.t.* $p_{n,i}$ *satisfies* $D_{n,i}^{(\xi_n)}(p_{n,i}) \triangleq \frac{\mathrm{d}\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})}{\mathrm{d}p_{n,i}} = d_{n,i}(p_{n,i}) - \xi_n\rho$.

3. *If* $D_{n,i}^{(\xi_n)}(0) \leq 0$, $\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})$ *is strictly decreasing in* $p_{n,i} \geq 0$*; otherwise,* $\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})$ *is strictly increasing in* $p_{k,n} \geq 0$ *till* $p_{n,i}^{(\xi_n)}$*, then is strictly decreasing in* $p_{n,i}$*, where* $D_{n,i}^{(\xi_n)}(p_{n,i}^{(\xi_n)}) = 0$.[3]

According to Theorem 6.3, $\sum_{i\in\mathcal{H}_n}\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})$ is strictly concave in $\mathbf{p}_n$. Since constraints in (6.6b) are affine, problem (6.6) is a concave optimization problem. The partial Lagrangian of problem (6.6) is

$$\mathcal{L}_n(\lambda_n,\mathbf{p}_n) = \sum_{i\in\mathcal{H}_n}\Big(C_{n,i}(p_{n,i}) - \xi_n\rho p_{n,i}\Big) - \lambda_n(\mathbf{1}_{|\mathcal{H}_n|}^T\cdot\mathbf{p}_n - p_n^{\max}),$$

where $\lambda_n \geq 0$ is the associated Lagrange multiplier. Then, the *Karush-Khun-Tucker* (KKT) conditions are

$$\lambda_n(\mathbf{1}_{|\mathcal{H}_n|}^T\cdot\mathbf{p}_n - p_n^{\max}) = 0,\ p_{n,i}\geq 0, \lambda_n \geq 0, \tag{6.10}$$

$$d_{n,i}(p_{n,i}) - \xi_n\rho - \lambda_n = 0,\ \text{if } p_{n,i} > 0. \tag{6.11}$$

For a given $\lambda_n \geq 0$, from (6.11) we have

$$p_{n,i} = \begin{cases} 0 & \text{if } d_{n,i}(0) \leq \xi_n\rho+\lambda_n \\ d_{n,i}^{-1}(\xi_n\rho+\lambda_n) & \text{if } d_{n,i}(0) > \xi_n\rho+\lambda_n, \end{cases} \tag{6.12}$$

where $d_{n,i}^{-1}(\cdot)$ denotes the inverse function of $d_{n,i}(\cdot)$. On the other hand, it is obvious $\lambda_n = d_{n,i}(p_{n,i}) - \xi_n\rho = D_{n,i}^{(\xi_n)}(p_{n,i})$ for subchannel i such that $p_{n,i} > 0$. Together with Theorem 6.3, it can be further inferred

$$\lambda_n \leq \max_{i\in\mathcal{H}_n} D_{n,i}^{(\xi_n)}(0). \tag{6.13}$$

Then, the unique optimal λ_n can be easily determined by bisection search in the interval $[0, \max_{i\in\mathcal{H}_n} D_{n,i}^{(\xi_n)}(0)]$. Specifically, if (6.10) holds for the $p_{n,i}$'s obtained in (6.12), then the current λ is the optimal one. Otherwise, if (6.10) does not hold and $\mathbf{1}_{|\mathcal{H}_n|}^T\mathbf{p}_n - p_n^{\max} > 0$, λ should

[3] $p_{n,i}^{(\xi_n)}$ must be finite since the strict concavity of $C_{n,i}(p_{n,i})$ enforces its derivative w.r.t $p_{n,i}$ strictly decreasing and eventually less than ξ_n. And $p_{n,i}^{(\xi_n)}$ can be efficiently solved using golden section search [96] or bisection search as a result of the concavity of $\mathcal{T}_{n,i}^{(\xi_n)}(p_{n,i})$.

Table 6.1. Energy-efficient Power Allocation for OFDMA-based two-way AF relay with Fixed Subchannel Assignment.

1. Initialization: $\xi_n \leftarrow 0$, $\lambda_n \leftarrow 0$, $\underline{\lambda}_n \leftarrow 0$, $\overline{\lambda}_n \leftarrow \max_{i\in\mathcal{H}_n} D_{n,i}^{(\xi_n)}(0)$
3. **while** no convergence for ξ_n
4. **while** no convergence for λ_n
5. Calculate $\mathbf{p}_n = (p_{n,i})_{i\in\mathcal{H}_n}$ using (6.12);
6. **if** $\mathbf{1}_{|\mathcal{H}_n|}^T \cdot \mathbf{p}_n = p_n^{\max}$ or $\lambda_n = 0$ && $\mathbf{1}_{|\mathcal{H}_n|}^T \cdot \mathbf{p}_n < p_n^{\max}$
7. **break**;
8. **else if** $\mathbf{1}_{|\mathcal{H}_n|}^T \cdot \mathbf{p}_n < p_n^{\max}$
9. $\overline{\lambda}_n \leftarrow \lambda_n$; % *the optimal* λ_n *should be smaller*
10. **else**
11. $\underline{\lambda}_n \leftarrow \lambda_n$; % *the optimal* λ_n *should be bigger*
12. **end if**
13. $\lambda \leftarrow \frac{1}{2}(\underline{\lambda}_n + \overline{\lambda}_n)$; % *bisection search*
14. **end while**
15. $\mathbf{p}_n^* \leftarrow \mathbf{p}_n$;
16. $\xi_n \leftarrow \frac{\sum_{i\in\mathcal{H}_n} C_{n,i}^*(p_{n,i}^*)}{4p_c^D|\mathcal{H}_n| + 4p_c^S + \rho\sum_{i\in\mathcal{H}_n} p_{n,i}^*}$;% *Newton's method*
17. **end while**
18. $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)} \leftarrow \xi_n$;

be increased; if (6.10) does not hold and $\mathbf{1}_{|\mathcal{H}_n|}^T \cdot \mathbf{p}_n - p_n^{\max} < 0$, λ should be decreased. Moreover, it is worth mentioning if $\sum_{i\in\mathcal{H}_n} \max(0, p_{n,i}^{(\xi_n)}) < p_n^{\max}$, the optimal λ_n is simply zero. In this case, the optimal transmit power of terminal pair n on the subchannels simply equal their respective $\max(0, p_{n,i}^{(\xi_n)})$. Note that this case is quite likely the most typical one for the energy-efficient resource allocation, considering that energy-efficient transmission usually does not operate at the full transmit power [83, 79, 90, 92]. The procedures for solving the subproblems for given subchannel assignment (i.e., problem (6.5)) are summarized in Table 6.1.

To obtain more insights about problem (6.5) such as the structure of the optimal solution, we further exploit the intrinsic properties of the problem. On the one hand, we can

obviously reformulate problem (6.5) as follows:

$$\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)} \equiv \max_{\mathbf{p}_n} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n), \tag{6.14a}$$

$$s.t. \quad (6.6b), \tag{6.14b}$$

where $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ denotes the maximum EE of terminal pair n given a transmit power vector, $\mathbf{p}_n$, and is defined as

$$\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n) \triangleq \frac{\sum_{i\in\mathcal{H}_n} C_{n,i}(p_{n,i})}{4p_c^D|\mathcal{H}_n| + 4p_c^S + \rho\sum_{i\in\mathcal{H}_n} p_{n,i}}.$$

Let $\mathbf{p}_n^{(\mathcal{H}_n)} \triangleq (p_{n,i}^{(\mathcal{H}_n)})_{i\in\mathcal{H}_n}$ represent one optimal $\mathbf{p}_n$ for problem (6.14). Then, we have the following theorem proved in Appendix E.3.

Theorem 6.4 *For any given $\mathcal{H}_n$, $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ is strictly pseudoconcave in $\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|}$. Therefore, there exits a unique $\mathbf{p}_n^{(\mathcal{H}_n)}$ that attains $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}$.*

Mathematically, pseudoconcavity is weaker than concavity but stronger than quasiconcavity [97]. Besides the uniqueness of global maximum, strict pseudoconcavity also implies that [97]: (i) a local maximum is the global maximum; (ii) a solution of the KKT optimality conditions is the global maximum. These nice properties potentially enable us to solve problem (6.5) using other approaches, e.g., gradient-based methods.

On the other hand, we can also reformulate problems (6.5) and (6.14) as follows:

$$\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)} \equiv \max_{0\le p_n\le p_n^{\max}} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(p_n), \tag{6.15a}$$

where $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(p_n)$ denotes the maximum EE of terminal pair n given a total transmit power, $p_n \triangleq \mathbf{1}_{2I}^T \cdot \overline{\mathbf{p}}_n$, and is defined as

$$\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(p_n) \triangleq \frac{C_n^{(\mathcal{H}_n)}(p_n)}{4p_c^D|\mathcal{H}_n| + 4p_c^S + \rho p_n},$$

while $C_n^{(\mathcal{H}_n)}(p_n)$ denotes the maximum exchanged rate of terminal pair n given p_n, and is a

concave function of p_n as follows:

$$C_n^{(\mathcal{H}_n)}(p_n) \triangleq \max_{\mathbf{p}_n} \sum_{i\in\mathcal{H}_n} C_{n,i}(p_{n,i}), \tag{6.16a}$$

$$s.t.\ \ \mathbf{1}_{|\mathcal{H}_n|}^T \cdot \mathbf{p}_n = p_n,\ \mathbf{p}_n \geq \mathbf{0}_{|\mathcal{H}_n|}. \tag{6.16b}$$

Let $p_n^{(\mathcal{H}_n)}$ represent one optimal p_n for problem (6.15).[4] Note that if $C_n^{(\mathcal{H}_n)}(p_n^{\max}) < r_n^{\min}$, problem (6.5) is infeasible. Then, we have the following theorem, which can be similarly proved as the last one.

Theorem 6.5 *For any given $\mathcal{H}_n$, $\mathcal{E}_n^{(\mathcal{H}_n)}(p_n)$ is strictly pseudoconcave in $p_n \geq 0$. Therefore, there exits a unique $p_n^{(\mathcal{H}_n)}$ that attains $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}$.*

This theorem indicates the $\mathcal{E}_n^{(\mathcal{H}_n)}(p_n)$-versus-$p_n$ curve appears like a bell shape, noticing that $\mathcal{E}_n^{(\mathcal{H}_n)}(p_n)$ approaches zero when p_n goes to infinity. That is, $\mathcal{E}_n^{(\mathcal{H}_n)}(p_n)$ first increases with p_n till attaining the maximum, then deceases with p_n; however, $C_n^{(\mathcal{H}_n)}(p_n)$ simply increases with p_n. This phenomenon implies the possible tradeoff relationship between throughput and EE for each terminal pair and validates the necessity of studying the energy-efficient OFDMA-based two-way AF relay.

6.3.2 Low-Complexity Subchannel Allocation

Before introducing the proposed low-complexity subchannel allocation approach, we consider the assignment of subchannels for terminal pair n ($n \in \mathcal{N}$) in a sequential manner. Let $i_n^{(k)}$ denote the kth assigned subchannel for terminal pair n. Denote $\mathcal{H}_n^{(k)}$ to be the set of the first k assigned subchannels for terminal pair n, i.e, $\mathcal{H}_n^{(k)} = \{i_n^{(1)}, i_n^{(2)}, \cdots, i_n^{(k)}\}$. Then, $\mathcal{H}_n^{(k)} = \mathcal{H}_n^{(k-1)} \bigcup \{i_n^{(k)}\}$ for $k \geq 1$, where $\mathcal{H}_n^{(0)} = \varnothing$. Denote $\mathcal{I}_n^{(k)}$ as the set of candidate subchannels for the kth subchannel assignment for terminal pair n. Without loss of generality, we make three assumptions about $\{\mathcal{I}_n^{(k)} | k = 1, \cdots\}$'s for sequential subchannel assignment:

1. The initial candidate subchannel sets for each terminal pair is identical to the set of all subchannels, i.e., $\mathcal{I}_n^{(1)} = \mathcal{I}$ for $n \in \mathcal{N}$.

[4] According to Theorem 6.5, the $p_n^{(\mathcal{H}_n)}$ defined here is identical to the one defined in Theorem 6.4.

2. The set of candidate subchannels for later subchannel assignment is a proper subset of that for earlier subchannel assignment, i.e., $\mathcal{I}_n^{(k+1)} \subsetneq \mathcal{I}_n^{(k)}$ for $k \geq 1$.

3. The chosen subchannel is removed for future subchannel assignment, i.e., $i_n^{(k)} \in \mathcal{I}_n^{(k)} \setminus \mathcal{I}_n^{(k+1)}$ for $k \geq 1$.[5]

The last two assumptions are based on the previous assumption that one subchannel can be assigned to at most one terminal pair.

Next, we define an EE-oriented sequential subchannel assignment policy that will facilitate the low-complexity subchannel allocation.

Definition 1 *Let $\mathcal{P}_n$ represent the EE-oriented sequential subchannel allocation policy for terminal pair n as follows: Given $\mathcal{H}_n^{(k)}$ and $\mathcal{I}_n^{(k+1)}$, $\mathcal{P}_n$ determines $i_n^{(k+1)}$ as*

$$i_n^{(k+1)} = \arg\max_{i \in \mathcal{I}_n^{(k+1)}} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k\star i)})}.$$

where $\mathcal{H}_n^{(k\star i)} \triangleq \mathcal{H}_n^{(k)} \bigcup \{i\}$.

Accordingly, $\mathcal{P}_n$ chooses the subchannel that helps get the maximum EE. Furthermore, the following two properties hold for $\mathcal{P}_n$.

Property 6.1 *If $\mathcal{P}_n$ is adopted as the subchannel assignment policy for terminal pair n, $C_n^{(\mathcal{H}_n^{(k)})}(p_n)$ increases (not necessarily strictly) with $k \geq 1$.*

Property 6.2 *If $\mathcal{P}_n$ is adopted as the subchannel assignment policy for terminal pair n, $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})}$ does not necessarily increase with $k \geq 1$.*

Property 6.1 is due to the fact that every additional subchannel enables a more spectral-efficient allocation of the same amount of transmit power. As for Property 6.2, if p_c^D is set zero in the circuit power model, the energy efficiency of terminal pair n simply increases (not necessarily strictly) with the number of assigned subchannels, $k \geq 1$, because of the

[5]Note that $i_n^{(k)}$ is not necessarily the only element of $\mathcal{I}_n^{(k)} \setminus \mathcal{I}_n^{(k+1)}$, considering that some elements of $\mathcal{I}_n^{(k)} \setminus \mathcal{I}_n^{(k+1)}$ may be assigned to other terminal pairs.

same reason as Property 6.1. However, the energy efficiency of terminal pair n may not necessarily increase with $k \geq 1$ when $p_c^D > 0$. This is because every additional subchannel incurs an increment of circuit power of p_c^D, counteracting the rate improvement from the additional subchannel. The uncertainty of the monotonicity of $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})}$ in $k \geq 1$ makes it difficult to decide whether to assign terminal pair n more subchannels or simply terminate, if we find $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})} > \widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k+1)})}$ but still look forward to the possibility of $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k')})} > \widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})}$ for a certain $k' > k+1$. Fortunately, the following theorem proved in Appendix E.4 and its derivative corollary can relieve us from the dilemma.

Theorem 6.6 *If $\mathcal{P}_n$ is adopted as the subchannel assignment policy for terminal pair n and the following condition: $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})} > \widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k+1)})}$ holds, then we have $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})} > \widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k')})}$, for $k' \geq k+2$.*

From Theorem 6.6, an early termination criterion for the EE-oriented sequential subchannel assignment policy, $\mathcal{P}_n$, without jeopardizing the optimality of the EE of terminal pair n is immediate as in the following corollary.

Corollary 6.1 *If $\mathcal{P}_n$ is adopted as the subchannel assignment policy for terminal pair n, the sequential subchannel assignment for terminal pair n can be terminated as soon as the condition $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})} > \widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k+1)})}$ is satisfied for a certain $k \geq 1$. And the maximum EE for terminal pair n by the sequential subchannel assignment is $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k^{(n)})})} = \max_{k \geq 1} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k^{(n)})})}$, where $k^{(n)}$ denotes the smallest k satisfying that condition.*

Based on the discussed EE-oriented sequential subchannel assignment policy with the sufficient condition for early termination, we propose a low-complexity suboptimal subchannel allocation approach for problem (6.4). The basic principle is to sequentially allocate the subchannels, aiming to gradually improve the aggregated EE utility among all terminals pairs till all subchannels are assigned or till none of the terminal pairs prefers more subchannels. Specifically, follow the EE-oriented sequential subchannel assignment policy, $\{\mathcal{P}_n | n \in \mathcal{N}\}$, at each round of the subchannel assignment to find the most favorite subchannel (i.e., the subchannel incurring highest EE increment) among the unassigned

ones for every terminal pair. Then, only the terminal pair with the highest nonnegative EE utility increment if added its most favorite subchannel is assigned one subchannel this round, i.e, the most favorite one. In addition, any terminal pair, whose EE utility increment is negative with its favorite subchannel, can be safely excluded from future subchannel assignment according to Corollary 1. The pseudocode of the proposed subchannel assignment approach is given in Table 6.2.

Table 6.2. Low-Complexity Subchannel Allocation for the OFDMA-based two-way AF relay.

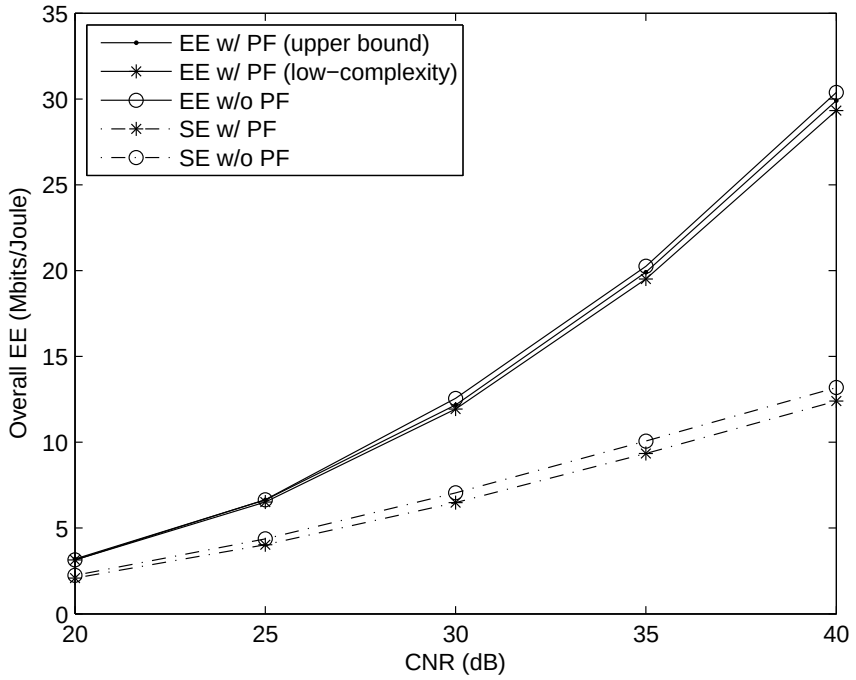
1. Initialization: $\mathcal{N}_+ \leftarrow \mathcal{N}$; $\widehat{\mathcal{E}_n} \leftarrow \delta, \mathcal{H}_n \leftarrow \varnothing, \forall n \in \mathcal{N}$,
   where δ is a relatively small positive number.
3. **while** $\mathcal{I} \neq \varnothing$ && $\mathcal{N}_+ \neq \varnothing$
4. **for** $n \in \mathcal{N}_+$
5. **for** $i \in \mathcal{I}$ % *i.e.,* $\mathcal{I}_n^{(|\mathcal{H}_n|+1)} = \mathcal{I}$
6. Calculate $\widehat{\mathcal{E}}_n^{(i)} \triangleq \widehat{\mathcal{E}}_n^{(\mathcal{H}_n \bigcup \{i\})}$ by solving problem (6.5);
7. $\delta_n^{(i)} \triangleq \ln \widehat{\mathcal{E}}_n^{(i)} - \ln \widehat{\mathcal{E}_n}$;
8. **end for**
9. **if** $\max\limits_{i \in \mathcal{I}} \delta_n^{(i)} < 0$
10. $\mathcal{N}_+ \leftarrow \mathcal{N}_+ \setminus \{n\}$; % *early termination by Corollary 1*
11. **else**
12. $i_n \leftarrow \arg\max\limits_{i \in \mathcal{I}} \delta_n^{(i)}$; % *by assignment Policy* $\mathcal{P}_n$
13. **end if**
14. **end for**
15. **if** $\mathcal{N}_+ = \varnothing$
16. **break**;
17. **end if**
18. $\bar{n} \leftarrow \max\limits_{n \in \mathcal{N}_+} \delta_n^{(i_n)}$; % *one with highest EE utility increment*
19. $\mathcal{H}_{\bar{n}} \leftarrow \mathcal{H}_{\bar{n}} \bigcup \{i_{\bar{n}}\}$; % *assigned a subchannel this round*
20. $\mathcal{I} \leftarrow \mathcal{I} \setminus \{i_{\bar{n}}\}$; % *implying* $\mathcal{I}_n^{(|\mathcal{H}_n|+1)} \subsetneq \mathcal{I}_n^{(|\mathcal{H}_n|)}$ *for* $n \in \mathcal{N}_+$
21. $\widehat{\mathcal{E}}_{\bar{n}} \leftarrow \widehat{\mathcal{E}}_n^{(i_{\bar{n}})}$;
22. **end while**

Next, we roughly analyze the complexity of the proposed suboptimal subchannel allocation algorithm. The major computational burden for each round of assignment comes from finding the most favorite subchannel and the corresponding EE utility increment for each terminal pair, i.e., solving problem (6.5). In practice, for terminal pair $n \in \mathcal{N}_+$ such that $n \neq \bar{n}$, Line 5-Line 13 in Table 6.2 do not need to be implemented at the new round since the required information is the same as that of the last round except for one subchannel is removed. The average total number of solving problem (6.5) is less than $NI + \sum_{k=2}^{I-1}(I-k) \leq NI + \frac{1}{2}I^2 \approx O\left(I^2\right)$.[6]

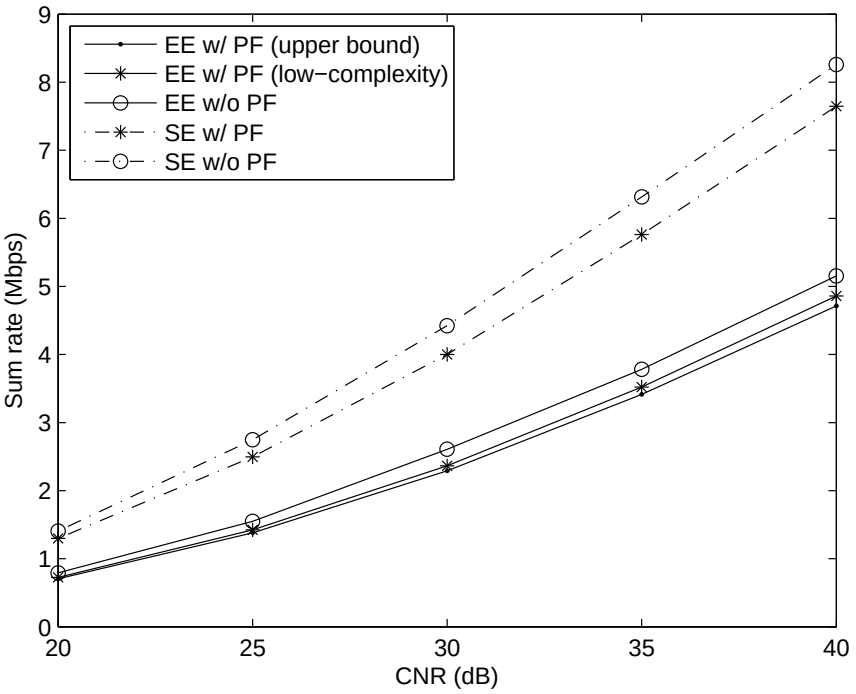
6.4 Numerical Results

In this section, we present simulation results to verify the benefits of the energy-efficient OFDMA-based two-way AF relay. There are $I = 40$ subchannels in total, each with a bandwith of 15 kHz. The maximum transmit power of the relay is $1\,W$ and its transmit power on any active subchannel is fixed at $q_i = 1/40\,\mathrm{W}$. For each terminal, the static circuit power is $p_c^S = 25\,\mathrm{mW}$; the dynamic circuit power for each active subchannel is $p_c^D = 0.4$ mW; and the power amplifier efficiency is $\frac{1}{\rho} = 40\%$. The maximum overall transmit power of each terminal pair is $p_n^{\max} = 200\,\mathrm{mW}$. For simplicity, we assume the average *channel-gain-to-noise ratios* (CNRs) (i.e., the large scale fading) for the links between the two terminals and the relay of a terminal pair are the same. The small scale fading for all links is assumed to be with Rayleigh distribution of unit average power. We consider the scenario with four terminal pairs. The first three ones have the same average CNRs for their links between the relay while the average CNRs of the fourth one is 10 dB higher. Besides the proposed energy-efficient OFDMA-based two-way AF relay with proportional fairness in EE, we also include the performance of the energy-efficient scheme without proportional fairness in EE and the spectral-efficient schemes with and without proportional fairness in data rate for comparison.

[6]The worst case is all the I subchannels are assigned.



(a) EE versus CNR



(b) Throughput versus CNR

Figure 6.2. The aggregated EE and rate performance of the energy-efficient schemes with and without proportional fairness in EE and the spectral-efficient schemes with and without proportional fairness in rate.

Figure 6.2 illustrates the aggregated EE (not the aggregated EE utility) and rate of the energy-efficient OFDMA-based two-way AF relay with and without proportional fairness in EE and the spectral-efficient OFDMA-based two-way AF relay with and without proportional fairness in rate. The x-axis is the CNR of the terminals with lower average CNR. It can be seen that the energy-efficient schemes can achieve much larger EE than the spectral-efficient ones, especially at the high CNR regime. The EE improvement of the energy-efficient schemes is at the cost of some rate loss compared to the spectral-efficient ones. And the overall EE degrades if provisioning proportional fairness in EE and rate for the energy- and spectral-efficient schemes, respectively. In addition, the proposed low-complexity scheme has close performance compared to the upper bound method.

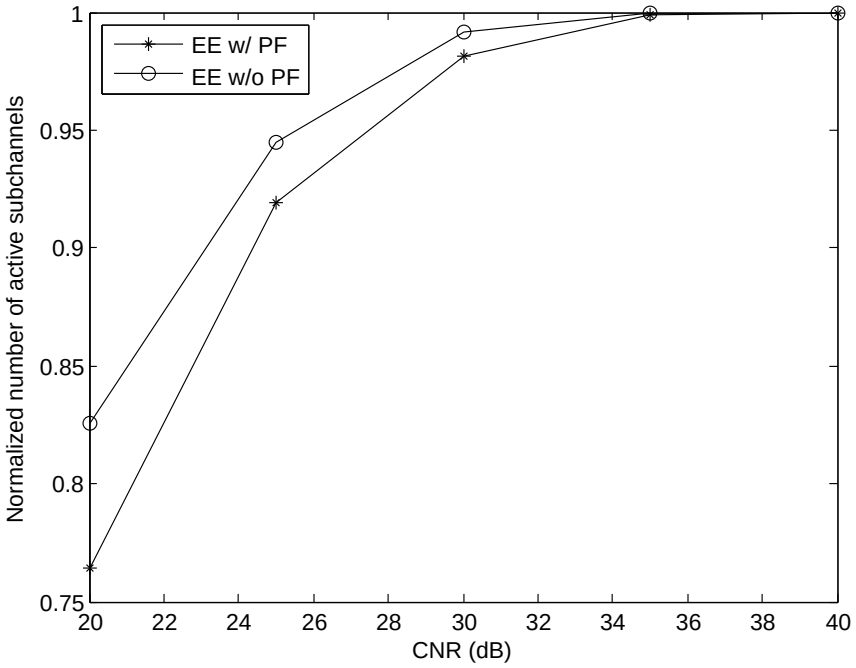


Figure 6.3. Normalized number of active subchannels. The x-axis is the CNR of the terminals with lower average CNR.

Figure 6.3 plots the normalized number of active subchannles. From it, it can be inferred that only a portion of the total subchannels are active at the low and medium CNR regimes for the energy-efficient schemes. This phenomenon indicates that the benefits of

adding more subchannels may not be worthy to justify the additional circuit power consumption for active subchannels for energy-efficient transmission. Moreover, it also implies the early termination of subchannel assignment, i.e., $\sum_{n\in\mathcal{N}} |\mathcal{H}_n| < N$, does happen. In the high CNR regime, all subchannels tend to be active; however, the total transmit powers of the energy-efficient schemes are still less than those of the spectral-efficient ones, considering the rate loss in Figure 6.2(b).

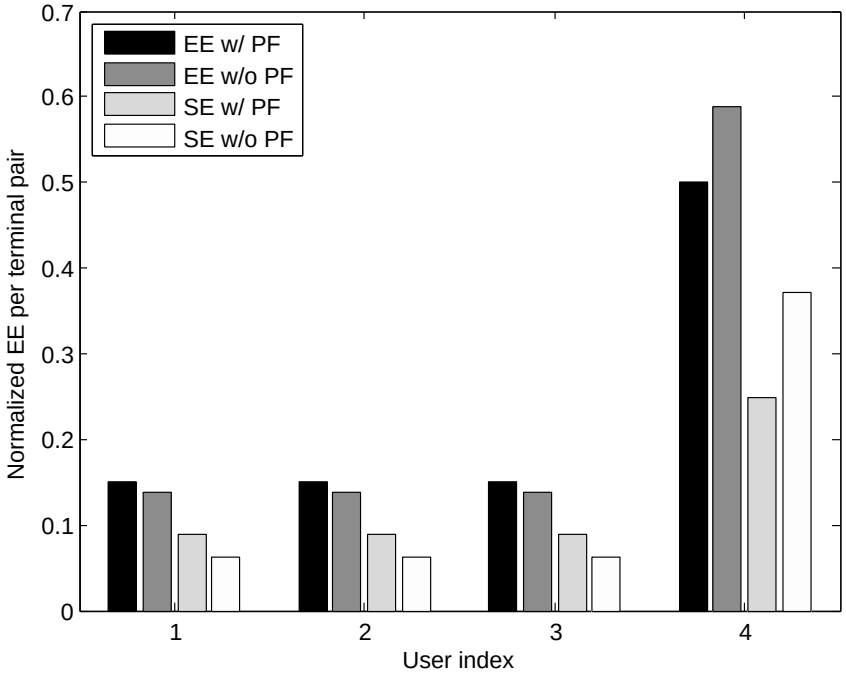


Figure 6.4. Normalized overall EE distribution among terminal pairs. The EEs are normalized by the overall EE of the energy-efficient scheme without proportional fairness in EE.

Finally, we compare the fairness in EE of the energy- and spectral-efficient schemes in Figure 6.4. It can be seen that the terminal pairs with lower CNRs can have larger EE if proportional fairness in EE is considered for the energy-efficient schemes and even if proportional fairness in rate is considered for the spectral-efficient scheme.

6.5 Conclusion

In this chapter, we have studied the energy-efficient resource allocation for OFDMA-based two-way relay, where joint power and subchannel allocation, including active subchannel

selection, is performed to reconcile the aggregated EE utility and the proportional fairness in EE among different terminal pairs. Without loss of generality, we set up a general EE optimization framework grounded on the AF relay, which can be extended to the DF and CF cases. The energy-efficient resource allocation problem is mixed-integer combinatorial and further proves to be nonconvex and NP-hard. In view of the difficulty in finding the global optimal, we first seek an upper-bound solution using continuous relaxation and the Lagrange dual method. To reduce the complexity, we exploit the the hidden concavity and the pseudoconcavity in the subproblems for any fixed subchannel assignment and propose an EE-oriented sequential subchannel assignment policy. Moreover, we discover the sufficient condition for the early termination of the sequential subchannel assignment without losing the EE optimality. Simulation results validate the great superiority in EE and proportional fairness of the energy-efficient OFDMA-based two-way relay over its spectral-efficient counterpart. Moreover, the proposed suboptimal scheme can efficiently achieve a satisfying EE performance.

CHAPTER 7

CONCLUSION

In this dissertation, we have studied the fundamental interrelationship between EE an SE and investigate energy-efficient resource allocation for pure OFDMA networks. We have also investigated joint energy-efficient design of OFDMA and CR and two-way relay, respectively. The main contributions are summarized as follows.

We have studied the fundamental interrelationship between EE and SE in downlink OFDMA networks and analyzed the impacts of channel gain and circuit power on the EE-SE relationship. We have established a general EE-SE optimization framework, where the overall EE, SE and per-user QoS are all considered. Under this framework, we have found that EE is quasiconcave in SE and decreases with SE when SE is large enough. Moreover, we have studied the energy-efficient resource allocation in both downlink and uplink OFDMA networks. For each scenario, we have found the optimal energy-efficient resource allocation approach and developed low-complexity suboptimal algorithm by exploring the inherent structure of the objective function and the feature of energy-efficient design. For the downlink case, we have also obtained a computationally efficient and numerically tractable EE upper bound, which tends to be rather tight if the number subcarriers is large compared with that of UEs and is the foundation of a near-optimal solution relying on the quasiconcave relation between the modified EE and transmit power. In addition, we have studied the statistical delay driven energy-efficient design in downlink OFDMA networks for delay-sensitive traffic and compared it with the spectral-efficient design. We have formulated SE and EE optimization problems with the statistical delay provisioning based on the concept of the effective capacity. We have proved and exploited the intrinsic quasiconcavity of EE on transmit power, which implicates the existence of a unique global maximum of EE. We have demonstrated that the resultant upper bound is quite tight when

the number of subcarriers is larger than that of the users. We have also analyzed the EE-delay tradeoff, the relationship between spectral-efficient and energy-efficient designs, and the impact of system parameters, such as circuit power and delay exponents.

We have also studied two types of energy-efficient spectrum access problems, which respectively optimize worst EE and average EE, in an OFDMA-based CR network. We have utilized continuous and concave relaxation to near-optimally solve the worst-EE-based problem. We have also developed a suboptimal alternative to further reduce complexity. For the average-EE-based problem, we have converted it into an equivalent form and introduced a concave envelope based B&B approach to find its optimal solution. We have then exploited some underlying properties of the energy-efficient transmission to speed up the convergence of the B&B approach. We have also developed a low-complexity suboptimal approach. Moreover, we have studied the energy-efficient resource allocation for OFDMA-based two-way relay, where joint power and subchannel allocation, including active subchannel selection, is performed to reconcile the aggregated EE utility and the proportional fairness in EE among different terminal pairs. Without loss of generality, we have set up a general EE optimization framework grounded on the AF relay, which can be extended to the DF and CF cases. In view of the difficulty in finding the global optimal, we have seeked an upper-bound solution using continuous relaxation and the Lagrange dual method. To reduce the complexity, we have exploited the the hidden concavity and the pseudoconcavity in the subproblems for any fixed subchannel assignment and proposed an EE-oriented sequential subchannel assignment policy. In addition, we have discovered the sufficient condition for the early termination of the sequential subchannel assignment without losing the EE optimality.

APPENDIX A

PROOF FOR CHAPTER 2

A.1 Proof of Theorem 2.1

Denote the superlevel sets of $\eta_{EE}^{*}(\boldsymbol{R})$ as

$$\mathcal{S}_{\alpha} = \{\boldsymbol{R} \geq \left[\check{R}_k\right]_{K\times 1} \,|\, \eta_{EE}^{*}(\boldsymbol{R}) \geq \alpha\}.$$

According to [76], $\eta_{EE}^{*}(\boldsymbol{R})$ is strictly quasiconcave in $\boldsymbol{R}$ if $\mathcal{S}_{\alpha}$ is strictly convex for any real number α. When $\alpha < 0$, no points exist on the counter $\eta_{EE}^{*}(\boldsymbol{R}) = \alpha$. When $\alpha \geq 0$, $\mathcal{S}_{\alpha}$ is equivalent to $\mathcal{S}_{\alpha} = \{\boldsymbol{R} \geq \left[\check{R}_k\right]_{K\times 1} \,|\, \alpha\zeta P^{*}(\boldsymbol{R}) + \alpha P_s + (\alpha\xi - 1)R \leq 0\}$, where $P^{*}(\boldsymbol{R})$ is the minimum total transmit power needed for any rate vector $\boldsymbol{R}$ that satisfies all constraints but not necessarily including the peak power one in (2.6). From [64, 59, 58], it is known that $P^{*}(\boldsymbol{R})$ is strictly convex in $\boldsymbol{R}$, given a sufficiently large number of subcarriers. As a result, $\mathcal{S}_{\alpha}$ is also strictly convex. Hence, Theorem 1 follows.

A.2 Proof of Theorem 2.2

Let $\boldsymbol{R}_1^{*}$, $\boldsymbol{R}_2^{*}$ and $\boldsymbol{R}_3^{*}$ denote the optimal rate vectors corresponding to the overall throughput R_1, R_2, and R_3, respectively, and they also satisfy all constraints but not necessarily including the peak power one in (2.6). Without loss of generality, assume that $R_1 < R_2 < R_3$. Let $\boldsymbol{R}_2$ denote the rate vector as follows.

$$\begin{aligned} \boldsymbol{R}_2 &= \frac{R_3 - R_2}{R_3 - R1}\boldsymbol{R}_1^{*} + \frac{R_2 - R_1}{R_3 - R1}\boldsymbol{R}_3^{*} \\ &= \lambda\boldsymbol{R}_1^{*} + (1-\lambda)\,\boldsymbol{R}_3^{*}, \end{aligned}$$

where $\lambda = \frac{R_3 - R_2}{R_3 - R1}$ and $0 < \lambda < 1$. Obviously, $\boldsymbol{R}_2$ is also in the feasible region of (2.6) and its sum rate is R_2. From [64, 59, 58], it is known that $P^{*}(\boldsymbol{R})$ is strictly convex in $\boldsymbol{R}$, given a sufficiently large number of subcarriers. Thus, $P^{*}(\boldsymbol{R_2}) < \lambda P^{*}(\boldsymbol{R}_1^{*}) + (1-\lambda)\,P^{*}(\boldsymbol{R}_3^{*})$. Since $\boldsymbol{R}_2^{*}$ is the optimal rate vector among all the rate vectors with a summation of R_2, we

have $P^*(\boldsymbol{R}_2^*) \leq P^*(\boldsymbol{R}_2)$. Consequently, we have that $P^*(\boldsymbol{R}_2^*) < \lambda P^*(\boldsymbol{R}_1^*) + (1-\lambda) P^*(\boldsymbol{R}_3^*)$. Thus, for any given throughput $R\left(= B\eta_{SE}\right)$, the minimum transmit power $P^*(R) = P^*(\boldsymbol{R}^*)$, is strictly convex in R (and η_{SE}).

Denote the superlevel set of $\eta_{EE}^*\left(\eta_{SE}\right)$ as

$$\mathcal{S}_\beta = \{R \geq \check{R} \,|\, \eta_{EE}^*\left(\eta_{SE}\right) \geq \beta, \beta \in \mathbb{R}\}.$$

Similar to the proof of Theorem 1, when $\beta \geq 0$, $\mathcal{S}_\beta$ is equivalent to $\mathcal{S}_\beta = \{R \geq \check{R} \,|\, \beta\zeta P^*\left(\eta_{SE}\right) + \beta P_s + B(\beta\xi - 1)\eta_{SE} \leq 0\}$, where $P^*\left(\eta_{SE}\right)$ is the minimum total transmit power needed for any SE $\eta_{SE} \geq \check{R}/B$. As a result of the convexity of $P^*\left(\eta_{SE}\right)$ (i.e., $P^*(R)$) proved above, $\mathcal{S}_\beta$ is strictly convex in η_{SE}. Hence, $\eta_{EE}^*\left(\eta_{SE}\right)$ is strictly quasiconcave and has a unique global maximum.

On the other hand, it is obvious that $\lim\limits_{\eta_{SE}\to\infty} \eta_{EE}^*\left(\eta_{SE}\right) = \lim\limits_{\eta_{SE}\to\infty} \max\limits_{\boldsymbol{\rho}\in\varrho, p_{k,n}\geq 0} \frac{B\eta_{SE}}{\zeta P^*\left(\eta_{SE}\right) + P_s + \xi B\eta_{SE}} = \lim\limits_{P^*\left(\eta_{SE}\right)\to\infty} \frac{o(P^*\left(\eta_{SE}\right))}{P^*\left(\eta_{SE}\right)} = 0$. Thus, starting from $\eta_{SE} = \frac{\check{R}}{B}$, $\eta_{EE}^*\left(\eta_{SE}\right)$ either strictly decreases with η_{SE} if $\left.\frac{d\eta_{EE}^*\left(\eta_{SE}\right)}{d\eta_{SE}}\right|_{\eta_{SE}=\check{R}/B} \leq 0$, or first strictly increases and then strictly decreases with η_{SE} if $\left.\frac{d\eta_{EE}^*\left(\eta_{SE}\right)}{d\eta_{SE}}\right|_{\eta_{SE}=\check{R}/B} > 0$. And the maximum EE in the SE region $\left[\frac{\check{R}}{B}, \frac{\hat{R}}{B}\right]$ is straightforward as indicated in Theorem 2.2. This completes the proof.

A.3 Proof of Properties 2.1 and 2.2

A.3.1 Proof of Property 2.1

$\eta_{EE}^*\left(\eta_{SE}\right) = \max\limits_{\boldsymbol{\rho}\in\varrho}\left(\max\limits_{p_{k,n}\geq 0} \eta_{EE}\left(\eta_{SE}, \boldsymbol{\rho}\right)\right) = \max\limits_{\boldsymbol{\rho}\in\varrho}\left(\frac{B\eta_{SE}}{P_c + \min\limits_{p_{k,n}\geq 0} \zeta P\left(\eta_{SE}, \boldsymbol{\rho}\right)}\right)$. For any given subcarrier allocation matrix $\boldsymbol{\rho} \in \varrho$, $P\left(\eta_{SE}, \boldsymbol{\rho}\right)$ can be minimized when power is distributed in a two-stage-water-filling fashion as stated in Theorem 2.3 and proved in Appendix D. Consequently, if any channel power gain $g_{k,n}$ increases, $\min\limits_{p_{k,n}\geq 0} P\left(\eta_{SE}, \boldsymbol{\rho}\right)$ strictly decreases or remains the same. Then, $\max\limits_{p_{k,n}\geq 0} \eta_{EE}\left(\eta_{SE}, \boldsymbol{\rho}\right)$ increases (but not necessarily strictly) with $g_{k,n}$. Thus, $\eta_{EE}^*\left(\eta_{SE}\right)$ increases with $g_{k,n}$, which means the $\eta_{EE}^*\left(\eta_{SE}\right)$-versus-$\eta_{SE}$ curve tends to be strictly higher or remains the same with the increase of $g_{k,n}$. Denote the new (with higher $g_{k,n}$) $\eta_{EE}^*\left(\eta_{SE}\right)$, η_{SE}^{opt} and $\eta_{SE}^{\max}$ as $\tilde{\eta}_{EE}^*\left(\eta_{SE}\right)$, $\tilde{\eta}_{SE}^{opt}$ and $\tilde{\eta}_{SE}^{\max}$, respectively. Then,

$\eta_{EE}^{opt} \leq \tilde{\eta}_{EE}^{*}\left(\eta_{SE}^{opt}\right) \leq \tilde{\eta}_{EE}^{opt}$ and $\eta_{EE}^{\max} \leq \tilde{\eta}_{EE}^{*}\left(\eta_{SE}^{\max}\right) \leq \tilde{\eta}_{EE}^{\max}$. Therefore, both the optimal EE, η_{EE}^{opt}, and the maximum EE, $\eta_{EE}^{\max}$, increase with channel power gain $g_{k,n}$.

A.3.2 Proof of Property 2.2

Since $\eta_{EE}^{*}\left(\eta_{SE}\right) = \frac{B\eta_{SE}}{\zeta P^{*}\left(\eta_{SE}\right)+P_c}$, $\eta_{EE}^{*}\left(\eta_{SE}\right)$ strictly decreases with P_c, which means the $\eta_{EE}^{*}\left(\eta_{SE}\right)$-versus-$\eta_{SE}$ curve tends to be strictly lower with the increase of P_c. Denote the new (with larger circuit power) $\eta_{EE}^{*}\left(\eta_{SE}\right)$, η_{SE}^{opt}, $\eta_{SE}^{\max}$ and P_c as $\tilde{\eta}_{EE}^{*}\left(\eta_{SE}\right)$, $\tilde{\eta}_{SE}^{opt}$, $\tilde{\eta}_{SE}^{\max}$, and $\tilde{P}_c$ respectively. Then, $\tilde{\eta}_{EE}^{opt} \leq \eta_{EE}^{*}\left(\tilde{\eta}_{SE}^{opt}\right) \leq \eta_{EE}^{opt}$ and $\tilde{\eta}_{EE}^{\max} \leq \eta_{EE}^{*}\left(\tilde{\eta}_{SE}^{\max}\right) \leq \eta_{EE}^{\max}$. Thus, η_{EE}^{opt} and $\eta_{EE}^{\max}$ strictly decrease with circuit power P_c.

On the other hand, in Case (i) of Theorem 2.2, $\eta_{SE}^{opt} = \eta_{SE}^{\max} = \frac{\check{R}}{B}$, thus it is nature that $\tilde{\eta}_{SE}^{opt} \geq \eta_{SE}^{opt}$ and $\tilde{\eta}_{SE}^{\max} \geq \eta_{SE}^{\max}$ as a result of the minimum throughput constraint. Furthermore, $\left.\frac{d\eta_{EE}^{*}(\eta_{SE})}{d\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{\max}} \leq 0$. Hence,

$$P^{*}(\eta_{SE}^{\max}) - \eta_{SE}^{\max} \left.\frac{dP^{*}\left(\eta_{SE}\right)}{d\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{\max}} + P_s \leq 0.$$

Thus, $\eta_{SE}^{\max}$ and η_{SE}^{opt} are both independent of ξ. In Cases (ii) and (iii), it is obvious that $\left.\frac{d\eta_{EE}^{*}(\eta_{SE})}{d\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{\max}} = 0$ and $\left.\frac{d\tilde{\eta}_{EE}^{*}(\eta_{SE})}{d\eta_{SE}}\right|_{\eta_{SE}=\tilde{\eta}_{SE}^{\max}} = 0$. Hence,

$$\eta_{SE}^{\max} \left.\frac{dP^{*}\left(\eta_{SE}\right)}{d\eta_{SE}}\right|_{\eta_{SE}=\eta_{SE}^{\max}} - P^{*}(\eta_{SE}^{\max}) = P_s,$$

$$\tilde{\eta}_{SE}^{\max} \left.\frac{dP^{*}\left(\eta_{SE}\right)}{d\eta_{SE}}\right|_{\eta_{SE}=\tilde{\eta}_{SE}^{\max}} - P^{*}(\tilde{\eta}_{SE}^{\max}) = \tilde{P}_s.$$

Since $\eta_{SE}\frac{dP^{*}(\eta_{SE})}{d\eta_{SE}} - P^{*}\left(\eta_{SE}\right)$ strictly increases with η_{SE} and $P_s < \tilde{P}_s$, we have $\eta_{SE}^{\max} < \tilde{\eta}_{SE}^{\max}$. Then, $\tilde{\eta}_{EE}^{*}\left(\eta_{SE}\right)$ increases with η_{SE} in the region $[\eta_{SE}^{opt}, \tilde{\eta}_{SE}^{\max}]$. Hence, $\tilde{\eta}_{EE}^{*}\left(\eta_{SE}^{opt}\right) \leq \tilde{\eta}_{EE}^{opt} = \tilde{\eta}_{EE}^{*}\left(\tilde{\eta}_{SE}^{opt}\right)$. Accordingly, we have $\eta_{SE}^{opt} \leq \tilde{\eta}_{SE}^{opt}$. Thus, η_{SE}^{opt} and $\eta_{SE}^{\max}$ both increase with P_s but have nothing to do with ξ. Now the proof of Property 2.2 completes.

A.4 Proof of Theorem 2.3

Before proving Theorem **??**, we first introduce a lemma as follows.

Lemma A.1 *Let* $\boldsymbol{P} = \vec{f}(R, \mathcal{S})$ *represents the power needed on each subcarrier by water-filling to fulfill* R *over* $\mathcal{S}$*. For any* $\Delta R \geq 0$*,* $\vec{f}(R + \Delta R, \mathcal{S}) = \vec{f}(R, \mathcal{S}) + \boldsymbol{P}(R, \Delta R)$*, where* $\boldsymbol{P}(R, \Delta R) \geq \mathbf{0}$ *is a non-negative power vector. (The proof is omitted.)*

Lemma 1 indicates that each dimension of $\vec{f}(R, \mathcal{S})$ is continuous and non-decreasing with R. Thus, a water-filling process can be generally regard as a series of successive sub-water-filling processes, where the latter starts filling water from the initial water level obtained by the former.

Subsequently, we give a brief proof of Theorem 2.3.

Let P_k^* be the minimum power for user k of (2.7) given a fixed subcarrier allocation strategy $\boldsymbol{\rho}$. First, if $\sum_{k=1}^{K} \check{R}_k = R$, it is obvious that P_k^*'s can be achieved by theorem 1. Note that only the first stage water-filling is needed.

Second, we consider the case that $\sum_{k=1}^{K} \check{R}_k < R$. For each user, its power should be in a water-filling fashion on its all available subcarriers to guarantee the optimality. That is, $P_k^* = \vec{f}\left(R_k^*, \mathcal{S}_k\right)$, where R_k^* is the corresponding rate for user k. According to Lemma 1, we have $\vec{f}\left(R_k^*, \mathcal{S}_k\right) = \vec{f}\left(\check{R}_k, \mathcal{S}_k\right) + \boldsymbol{P}\left(\check{R}_k, R_k^* - \check{R}_k\right)$. Thus, it indicates that the first stage water-filling in Theorem 1 does not bring any loss of optimality yet. Further, with all the individual minimum rate constraints being satisfied after the first stage processing, it is better to distribute $\left(P_k^* - f\left(\check{R}_k, \mathcal{S}_k\right)\right)$'s among all the N subcarriers with water-filling than distribute them among each user's own available subcarriers with water-filling to satisfy the total rate. Since P_k^*'s are already optimal as assumed, the second phase processing as shown in Theorem 2.3 should derive exactly the same results. Now Theorem 2.3 follows.

A.5 Proof of Theorem 2.4

As shown in the proof of Theorem 2.2, with sufficiently many subcarriers, $P^*(R)$ is strcitly convex in R. Thus, $\eta_{EE}^*(\eta_{SE})$ is not only quasiconcave but also continuously differentiable in R. Since $\lim_{\Delta\eta_{SE}\to 0} \frac{\eta_{EE}^*(\eta_{SE}+\Delta\eta_{SE}) - \eta_{EE}^*(\eta_{SE})}{\Delta\eta_{SE}} = \lim_{\Delta R\to 0} B \frac{1 - \frac{\zeta\Delta P^*(R) + \xi\Delta R}{\Delta R}\eta_{EE}^*(\eta_{SE})}{\zeta(P^*(R)+\Delta P^*(R)) + \xi(R+\Delta R) + P_s}$, $\operatorname{sgn}\left(\frac{d\eta_{EE}^*(\eta_{SE})}{d\eta_{SE}}\right) = \operatorname{sgn}\left(\frac{1}{\zeta\frac{dP^*(R)}{dR} + \xi} - \eta_{EE}^*(\eta_{SE})\right)$, where $\operatorname{sgn}(x)$ denotes the sign of x. Let $\sum_{k\in\mathcal{K}_2}\sum_{n\in\mathcal{S}_k} \Delta r_{k,n}^* = \Delta R$.

Then,

$$\begin{aligned}
\frac{dP^*(R)}{dR} &= \lim_{\Delta R\to 0} \frac{\min\limits_{\Delta r^*_{k,n}\geq 0} \sum\limits_{k\in\mathcal{K}_2, n\in\mathcal{S}^*_k} \frac{N_0 W}{g_{k,n}}\left(2^{\frac{r^*_{k,n}+\Delta r^*_{k,n}}{W}} - 2^{\frac{r^*_{k,n}}{W}}\right)}{\Delta R} \\
&= \lim_{\Delta R\to 0} \frac{\min\limits_{\Delta r^*_{k,n}\geq 0} \sum\limits_{k\in\mathcal{K}_2, n\in\mathcal{S}_k} \frac{N_0 W}{g_{k,n}} 2^{\frac{r^*_{k,n}}{W}} \frac{\Delta r^*_{k,n}}{W} \ln 2}{\Delta R} \\
&= \min_{k\in\mathcal{K}_2, n\in\mathcal{S}_k} \frac{2^{\frac{r^*_{k,n}}{W}}}{g_{k,n}} N_0 \ln 2 \\
&= \min_{k\in\mathcal{K}_2, n\in\mathcal{S}_k} \frac{\ln 2}{W}\left(p^*_{k,n} + \frac{N_0 W}{g_{k,n}}\right).
\end{aligned}$$

A.6 Proof of Theorem 2.5

Let $\mathcal{B}_k\,(k = 1, 2, \cdots, K)$ be the set of subcarriers assigned to user k derived by the TPLB Algorithm, and $|\mathcal{B}_k| = l_k$. $P_{lb} = \sum_{k=1}^{K} f(R_k, \mathcal{B}_k)$. Let $\mathcal{S}^*_k$ be the set of subcarriers assigned to user k of the optimal solution to (2.15), and $|\mathcal{S}^*_k| = m^*_k$. $P^* = \sum_{k=1}^{K} f\left(r_k, \mathcal{S}^*_k\right)$. Let $\hat{\mathcal{S}}^*_k$ be composed of the m^*_k best subcarriers of user k among all the N subcarriers, and $\check{P}^* = \sum_{k=1}^{K} f\left(R_k, \hat{\mathcal{S}}^*_k\right)$. Obviously, we have that $\check{P}^* \leq P^*$. Next, we prove that $P_{lb} \leq \check{P}^*$.

If $l_k = m^*_k$ for $\forall\, k$, then $P_{lb} = \check{P}^* \leq P^*$. Otherwise, $\exists i, j$ such that $l_i > m^*_i$ and $l_j < m^*_j$. In this case, we first prove that

$$f(R_i, \mathcal{B}_i) + f\left(R_j, \mathcal{B}_j\right) \leq f\left(R_i, \mathcal{B}_i^{(l_i-1)}\right) + f\left(R_j, \mathcal{B}_j^{(l_j+1)}\right), \qquad (*)$$

where $\mathcal{B}_i^{(l_i-1)}$ is composed of the $l_i - 1$ best subcarriers of user i and $\mathcal{B}_j^{(l_j+1)}$ is composed of the $l_j + 1$ best subcarriers of user j.

Without loss of generality, assume that the last $t\,(0 \leq t \leq l_j)$ subcarriers of $\mathcal{B}_j$ are added after the l_ith subcarrier of $\mathcal{B}_i$ is added during the process of TPLB algorithm. Obviously, $f\left(R_j, \mathcal{B}_j^{(l_j-t)}\right) - f\left(R_j, \mathcal{B}_j^{(l_j-t+1)}\right) \leq f\left(R_i, \mathcal{B}_i^{(l_i-1)}\right) - f(R_i, \mathcal{B}_i)$. And as a result of the convexity stated in Theorem 2.3, $f\left(R_j, \mathcal{B}_j\right) - f\left(R_j, \mathcal{B}_j^{(l_j+1)}\right) \leq f\left(R_j, \mathcal{B}_j^{(l_j-t)}\right) - f\left(R_j, \mathcal{B}_j^{(l_j-t+1)}\right)$. With the above two inequalities, we have $f\left(R_j, \mathcal{B}_j\right) - f\left(R_j, \mathcal{B}_j^{(l_j+1)}\right) \leq f\left(R_i, \mathcal{B}_i^{(l_i-1)}\right) - f(R_i, \mathcal{B}_i)$.

Thus, (*) holds. And (*) can be similarly extended as follows.

$$f(R_i, \mathcal{B}_i) + f\left(R_j, \mathcal{B}_j\right) \leq f\left(R_i, \mathcal{B}_i^{(l_i - s)}\right) + f\left(R_j, \mathcal{B}_j^{(l_j + s)}\right), \qquad (**)$$

where $1 \leq s \leq l_i$.

On the other hand, $\{\hat{\mathcal{S}}_k^*\}$ can be easily obtained from $\{\mathcal{B}_k\}$ by reducing the number of subcarriers of sets such that $l_i \geq m_i^*$ in backward order but adding the number of subcarriers such that $m_j^* \leq l_j$ in $\{\mathcal{B}_k\}$. And according to (**), during this process, the total power required increases. Therefore, $P_{lb} \leq \check{P}^* \leq P^*$.

In the case that each user suffers flat-fading, it is obvious that $P^* \leq P_{lb} \leq \check{P}^* = P^*$ since $\{\mathcal{B}_k\}$ is also a feasible solution. Hence, we have $P_{lb} = P^*$. This completes the proof of Theorem 2.5.

APPENDIX B

PROOF FOR CHAPTER 3

B.1 Proof of Theorem 3.1

We first prove that $\hat{R}_w^{(\boldsymbol{\rho})}(P)$ under the constraint (3.10b) is strictly concave and continuously differentiable in P. With the nature of water-filling, it is easy to prove that the transmit power on each subcarrier is nondecreasing with the total transmit power. Then we consider the limit under the constraint $\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{S}_k}\Delta p_{k,n}=\Delta P$ in (B.1), where $\overset{(a)}{=}$ follows the fact that $\ln(1+ax)=ax+o(x)$ when $x\to 0$ and we do have $\Delta p_{k,n}\to 0$ since $\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{S}_k}\Delta p_{k,n}=\Delta P\to 0$ and $\Delta p_{k,n}\geq 0$. The existence of the limit indicates that $\hat{R}_w^{(\boldsymbol{\rho})}(P)$ is continuously differentiable in P and $\frac{\mathrm{d}\hat{R}_w^{(\boldsymbol{\rho})}(P)}{\mathrm{d}P}=\max_{k\in\mathcal{K},n\in\mathcal{S}_k}\frac{\omega_k W g_{k,n}\log_2 e}{1+g_{k,n}\hat{p}_{k,n}}$. Accordingly, $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ is continuously differentiable in P. Moreover, $\frac{\omega_k W g_{k,n}\log_2 e}{1+g_{k,n}\hat{p}_{k,n}}$ is nonincreasing with P for $k\in\mathcal{K}$ and $n\in\mathcal{S}_k$ while $\max_{k\in\mathcal{K},n\in\mathcal{S}_k}\frac{\omega_k W g_{k,n}\log_2 e}{1+g_{k,n}\hat{p}_{k,n}}$ is strictly monotonically decreasing with P. Thus, $\frac{\mathrm{d}^2\hat{R}_w^{(\boldsymbol{\rho})}(P)}{\mathrm{d}P^2}<0$ and $\hat{R}_w^{(\boldsymbol{\rho})}(P)$ is strictly concave in P.

Denote the superlevel sets of $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ as

$$\mathcal{S}_\beta=\{P\geq\textstyle\sum_{k\in\mathcal{K}}R_k^{-1}\left(\mathcal{S}_k,\check{R}_k\right)\mid\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)\geq\beta\}.$$

According to [76], $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)$ is strictly quasiconcave in P if $\mathcal{S}_\beta$ is strictly convex for any real number β. When $\beta<0$, no points exist on the counter $\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)=\beta$. When $\beta\geq 0$, $\mathcal{S}_\beta$ is equivalent to $\mathcal{S}_\beta=\{P\geq\sum_{k\in\mathcal{K}}R_k^{-1}\left(\mathcal{S}_k,\check{R}_k\right)\mid\beta\zeta P+\beta P_c-\hat{R}_w^{(\boldsymbol{\rho})}(P)\leq 0\}$. Since $\hat{R}_w^{(\boldsymbol{\rho})}(P)$ is strictly concave in P, $\mathcal{S}_\beta$ is strictly convex in P. This completes the proof of Property (i).

Since $\frac{\mathrm{d}\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)}{\mathrm{d}P}=\lim_{\Delta P\to 0}\frac{\frac{\hat{R}_w^{(\boldsymbol{\rho})}(P+\Delta P)}{\zeta(P+\Delta P)+P_c}-\frac{\hat{R}_w^{(\boldsymbol{\rho})}(P)}{\zeta P+P_c}}{\Delta P}=\lim_{\Delta P\to 0}\frac{\frac{\hat{R}_w^{(\boldsymbol{\rho})}(P+\Delta P)-\hat{R}_w^{(\boldsymbol{\rho})}(P)}{\Delta P}-\zeta\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)}{\zeta(P+\Delta P)+P_c}=\frac{\frac{\mathrm{d}\hat{R}_w^{(\boldsymbol{\rho})}(P)}{\mathrm{d}P}-\zeta\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)}{\zeta P+P_c}$, $\operatorname{sgn}\left(\frac{\mathrm{d}\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)}{\mathrm{d}P}\right)=\operatorname{sgn}\left(\frac{\mathrm{d}\hat{R}_w^{(\boldsymbol{\rho})}(P)}{\mathrm{d}P}-\zeta\hat{\eta}_{EE}^{(\boldsymbol{\rho})}(P)\right)$, where $\operatorname{sgn}(x)$ denotes the sign of x. Accordingly, Property (iii) is proved.

$$\lim_{\Delta P\to 0}\frac{\hat{R}_w^{(\rho)}(P+\Delta P)-\hat{R}_w^{(\rho)}(P)}{\Delta P} = \lim_{\Delta P\to 0}\frac{\left\{\max\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{S}_k}\omega_k W\log_2\left(1+g_{k,n}(\hat{p}_{k,n}+\Delta p_{k,n})\right)\right\}-\sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{S}_k}\omega_k W\log_2\left(1+g_{k,n}\hat{p}_{k,n}\right)}{\Delta P}$$
$$= \lim_{\Delta P\to 0}\frac{\max\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{S}_k}\omega_k W\log_2\left(\frac{1+g_{k,n}(\hat{p}_{k,n}+\Delta p_{k,n})}{1+g_{k,n}\hat{p}_{k,n}}\right)}{\Delta P} \tag{B.1}$$
$$\overset{(a)}{=} \lim_{\Delta P\to 0}\frac{\max\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{S}_k}\omega_k W\left[\frac{g_{k,n}\log_2 e}{1+g_{k,n}\hat{p}_{k,n}}\Delta p_{k,n}+o(\Delta p_{k,n})\right]}{\Delta P}$$
$$= \lim_{\Delta P\to 0}\frac{\max\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{S}_k}\frac{\omega_k W g_{k,n}\log_2 e}{1+g_{k,n}\hat{p}_{k,n}}\Delta p_{k,n}}{\Delta P} = \max_{k\in\mathcal{K},n\in\mathcal{S}_k}\frac{\omega_k W g_{k,n}\log_2 e}{1+g_{k,n}\hat{p}_{k,n}}.$$

On the other hand, with the strict concavity of $\hat{R}_w^{(\rho)}(P)$, it is obvious that $\lim\limits_{P\to\infty}\hat{\eta}_{EE}^{(\rho)}(P) = \lim\limits_{P\to\infty}\frac{\hat{R}_w^{(\rho)}(P)}{\zeta P+P_c} = \lim\limits_{P\to\infty}\frac{o(P)}{\zeta P+P_c} = 0$. Thus, starting from $P_0 = \sum_{k\in\mathcal{K}} R_k^{-1}\left(\mathcal{S}_k,\check{R}_k\right)$, $\hat{\eta}_{EE}^{(\rho)}(P)$ either strictly decreases with P if $\frac{\mathrm{d}\hat{\eta}_{EE}^{(\rho)}(P)}{\mathrm{d}P}|_{P=P_0}\leq 0$ or first strictly increases and then strictly decreases with P if $\frac{\mathrm{d}\hat{\eta}_{EE}^{(\rho)}(P)}{\mathrm{d}P}|_{P=P_0}>0$. This completes the proof of Property (ii).

Next, we derive (3.12) and (3.13). The basic idea of the power allocation process is firstly to allocate power to make each UE merely satisfy its rate requirement (Eq. (3.12)) then allocate the remaining power to the subcarriers that can further maximize the WSR (Eq. (3.13)). It can be straightforwardly realized by the method of Lagrange multiplier.

First, we solve the power minimization problem for each UE, that is, $\min_{p_{k,n}\geq 0}\sum_{n\in\mathcal{S}_k} p_{k,n}$ subject to $\sum_{n\in\mathcal{S}_k} r_{k,n} = \check{R}_k$. Its Lagrangian function can be expressed as $\mathcal{L}_k = \sum_{n\in\mathcal{S}_k} p_{k,n} - \lambda_k\left(\sum_{n\in\mathcal{S}_k} r_{k,n}-\check{R}_k\right)$, where λ_k is the Lagrange multiplier. Then, by differentiating $\mathcal{L}_k$ with respect to $p_{k,n}$'s and setting it zero, we have that

$$\frac{\partial\mathcal{L}_k}{\partial p_{k,n}} = 1-\frac{\lambda_k W}{\ln 2}\frac{g_{k,n}}{1+g_{k,n}p_{k,n}} = 0.$$

Then, $p_{k,n} = \lambda_k W\log_2 e-\frac{1}{g_{k,n}}$. Plugging it into (1), we have that $r_{k,n} = W\log_2\left(g_{k,n}\lambda_k W\log_2 e\right)$. Let $\mu_k \triangleq \lambda_k W\log_2 e$. Note that we require $p_{k,n}\geq 0$. Therefore, we have that $\check{p}_{k,n} = \left(\mu_k-\frac{1}{g_{k,n}}\right)^+$ and $\sum_{n\in\{n\in\mathcal{S}_k|\check{p}_{k,n}>0\}} W\log_2(\mu_k g_{k,n}) = \check{R}_k$.

Subsequently, we need to solve the WSR maximization problem, that is, $\max_{\Delta p_{k,n}\geq 0}\sum_{k\in\mathcal{K}}\omega_k\sum_{n\in\mathcal{S}_k} r_{k,n}$ subject to $p_{k,n} = \check{p}_{k,n}+\Delta p_{k,n}$ and $\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{S}_k}\Delta p_{k,n} = P-\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{S}_k}\check{p}_{k,n}$. Its Lagrangian

function can be constructed as $\mathcal{L} = \sum_{k\in\mathcal{K}} \omega_k \sum_{n\in\mathcal{S}_k} r_{k,n} - \nu \left(\sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{S}_k} \Delta p_{k,n} - \left(P - \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{S}_k} \check{p}_{k,n} \right) \right)$, where ν the Lagrange multiplier. Then, by differentiating $\mathcal{L}$ with respect to $\Delta p_{k,n}$'s and setting it zero, we have that

$$\frac{\partial \mathcal{L}}{\partial \Delta p_{k,n}} = \frac{\omega_k W}{\ln 2} \frac{g_{k,n}}{1 + g_{k,n}\left(\check{p}_{k,n} + \Delta p_{k,n}\right)} - \nu = 0.$$

Then, $\Delta p_{k,n} = \frac{\omega_k W}{\nu \ln 2} - \frac{1}{g_{k,n}} - \check{p}_{k,n}$. Let $\mu \triangleq \frac{W}{\nu \ln 2}$. Note that we require $\Delta p_{k,n} \geq 0$. Hence, we have that $\hat{p}_{k,n} = \check{p}_{k,n} + \Delta p_{k,n} = \check{p}_{k,n} + \left(\omega_k \mu - \frac{1}{g_{k,n}} - \check{p}_{k,n} \right)^+$ and $\sum_{k\in\mathcal{K}} \sum_{n\in\{n\in\mathcal{S}_k | \Delta p_{k,n} > 0\}} \left(\omega_k \mu - \frac{1}{g_{k,n}} - \check{p}_{k,n} \right) = P - \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{S}_k} \check{p}_{k,n}$.

B.2 Proof of Theorem 3.2

Define $\hat{\mathcal{R}}_w(\boldsymbol{P}) \triangleq \max_{\tilde{\boldsymbol{\rho}}\in\tilde{\mathcal{Q}}} \mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P})$ as the maximum MWSR under the given power allocation matrix, $\boldsymbol{P}$, and constraint (3.17b). In [70], $\mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P})$ without constraint (3.17b) is proved to be strictly and jointly concave in $\tilde{\rho}_{k,n}$ and $p_{k,n}$. Clearly, $\sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n} \tilde{r}_{k,n}$ can be expressed in the form of $\mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P})$ by letting $\omega_k = 1$ and $\omega_{k'} = 0$ for $k' \neq k$. Since $\mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P})$ without constraint (3.17b) is strictly and jointly concave in $\tilde{\rho}_{k,n}$ and $p_{k,n}$ for any given ω_k's, $-\sum_{n\in\mathcal{N}} \tilde{\rho}_{k,n} \tilde{r}_{k,n} \leq -\check{R}_k$ in (3.17b) naturally turns out to be a convex constraint. Hence, $\mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \boldsymbol{P})$ with constraint (3.17b) is also strictly and jointly concave in $\tilde{\rho}_{k,n}$ and $p_{k,n}$. Thus, for $\tilde{\boldsymbol{\rho}}_1$, $\tilde{\boldsymbol{\rho}}_2$, and $\boldsymbol{P}_1 \neq \boldsymbol{P}_2$ that satisfy (3.17b), and $\theta \in (0, 1)$, we have that $\mathcal{R}_w(\theta\tilde{\boldsymbol{\rho}}_1 + (1-\theta)\tilde{\boldsymbol{\rho}}_2, \theta\boldsymbol{P}_1 + (1-\theta)\boldsymbol{P}_2) > \theta\mathcal{R}_w(\tilde{\boldsymbol{\rho}}_1, \boldsymbol{P}_1) + (1-\theta)\mathcal{R}_w(\tilde{\boldsymbol{\rho}}_2, \boldsymbol{P}_2)$. On the other hand, for $\epsilon > 0$, there exist $\tilde{\boldsymbol{\rho}}_1^o$ and $\tilde{\boldsymbol{\rho}}_2^o$ such that $\mathcal{R}_w(\tilde{\boldsymbol{\rho}}_i^o, \boldsymbol{P}_i) > \hat{\mathcal{R}}_w(\boldsymbol{P}_i) - \epsilon$ for $i = 1, 2$. Then, we have

$$\begin{aligned}
\hat{\mathcal{R}}_w(\theta\boldsymbol{P}_1 + (1-\theta)\boldsymbol{P}_2) &= \max_{\tilde{\boldsymbol{\rho}}\in\tilde{\mathcal{Q}}} \mathcal{R}_w(\tilde{\boldsymbol{\rho}}, \theta\boldsymbol{P}_1 + (1-\theta)\boldsymbol{P}_2) \\
&\geq \mathcal{R}_w(\theta\tilde{\boldsymbol{\rho}}_1^o + (1-\theta)\tilde{\boldsymbol{\rho}}_2^o, \theta\boldsymbol{P}_1 + (1-\theta)\boldsymbol{P}_2) \\
&> \theta\mathcal{R}_w(\tilde{\boldsymbol{\rho}}_1^o, \boldsymbol{P}_1) + (1-\theta)\mathcal{R}_w(\tilde{\boldsymbol{\rho}}_2^o, \boldsymbol{P}_2) \\
&> \theta\hat{\mathcal{R}}_w(\boldsymbol{P}_1) + (1-\theta)\hat{\mathcal{R}}_w(\boldsymbol{P}_2) - \epsilon. \qquad \text{(B.2)}
\end{aligned}$$

Since (B.2) holds for any $\epsilon > 0$, $\hat{\mathcal{R}}_w(\theta\boldsymbol{P}_1 + (1-\theta)\boldsymbol{P}_2) > \theta\hat{\mathcal{R}}_w(\boldsymbol{P}_1) + (1-\theta)\hat{\mathcal{R}}_w(\boldsymbol{P}_2)$. Hence, $\hat{\mathcal{R}}_w(\boldsymbol{P})$ is strictly concave in $p_{k,n}$.

Define $\hat{\mathcal{R}}_w(P) \triangleq \max_{p_{k,n}\geq 0} \hat{\mathcal{R}}_w(\boldsymbol{P})$ under constraints (3.17b) and (3.17c). Let $\boldsymbol{P}_j^*$ be the transmit power matrix corresponding to $\hat{\mathcal{R}}_w\left(P_j\right)$, where $j = 1, 2, 3$. Without loss of generality, assume that $P_1 < P_3 < P_2$. Let $\boldsymbol{P}_3 = \frac{P_2-P_3}{P_2-P_1}\boldsymbol{P}_1^* + \frac{P_3-P_1}{P_2-P_1}\boldsymbol{P}_2^* = \vartheta\boldsymbol{P}_1^* + (1-\vartheta)\boldsymbol{P}_2^*$, where $\vartheta \triangleq \frac{P_2-P_3}{P_2-P_1}$. Obviously, the sum of all entries of $\boldsymbol{P}_3$ equals P_3. Then, $\hat{\mathcal{R}}_w(P_3) = \hat{\mathcal{R}}_w(\boldsymbol{P}_3^*) \geq \hat{\mathcal{R}}_w(\boldsymbol{P}_3) > \vartheta\hat{\mathcal{R}}_w(\boldsymbol{P}_1^*) + (1-\vartheta)\hat{\mathcal{R}}_w(\boldsymbol{P}_2^*) = \vartheta\hat{\mathcal{R}}_w(P_1) + (1-\vartheta)\hat{\mathcal{R}}_w(P_2)$. Thus, $\hat{\mathcal{R}}_w(P)$ is strictly concave in P. Accordingly, the quasiconcavity of $\hat{\eta}_{EE}^{UB}(P)$ in P can be proved in the same way as Theorem 3.1.

Let $\tilde{\boldsymbol{\rho}}^* = [\tilde{\rho}_{k,n}^*]_{K\times N}$ and $\boldsymbol{P}^* = [\tilde{\rho}_{k,n}]_{K\times N}$ be the matrices corresponding to $\hat{\mathcal{R}}_w(P)$, i.e., $\hat{\mathcal{R}}_w(P) = \mathcal{R}_w(\tilde{\boldsymbol{\rho}}^*, \boldsymbol{P}^*)$. Similarly, let $\hat{\mathcal{R}}_w(P+\Delta P) = \mathcal{R}_w(\tilde{\boldsymbol{\rho}}^* + \Delta\tilde{\boldsymbol{\rho}}^*, \boldsymbol{P}^* + \Delta\boldsymbol{P}^*)$. Let $\hat{\mathcal{R}}_w^{(\tilde{\boldsymbol{\rho}}^*)}(P+\Delta P) \triangleq \max_{\Delta p_{k,n}\geq 0} \mathcal{R}_w(\tilde{\boldsymbol{\rho}}^*, \boldsymbol{P}^* + \Delta\boldsymbol{P})$ and $\hat{\mathcal{R}}_w^{(\tilde{\boldsymbol{\rho}}^*+\Delta\tilde{\boldsymbol{\rho}}^*)}(P) \triangleq \max_{\Delta p_{k,n}\geq 0} \mathcal{R}_w(\tilde{\boldsymbol{\rho}}^* + \Delta\tilde{\boldsymbol{\rho}}^*, \boldsymbol{P}^* + \Delta\boldsymbol{P}^* - \Delta\boldsymbol{P})$ with $\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}} \Delta p_{k,n} = \Delta P > 0$. Then we have that $\hat{\mathcal{R}}_w^{(\tilde{\boldsymbol{\rho}}^*)}(P + \Delta P) \leq \hat{\mathcal{R}}_w(P + \Delta P)$ and $\hat{\mathcal{R}}_w^{(\tilde{\boldsymbol{\rho}}^*+\Delta\tilde{\boldsymbol{\rho}}^*)}(P) \leq \hat{\mathcal{R}}_w(P)$. Thus, we have the inequalities in (B.3).

$$\frac{\hat{\mathcal{R}}_w^{(\tilde{\boldsymbol{\rho}}^*)}(P + \Delta P) - \mathcal{R}_w(\tilde{\boldsymbol{\rho}}^*, \boldsymbol{P}^*)}{\Delta P} \leq \frac{\hat{\mathcal{R}}_w(P + \Delta P) - \hat{\mathcal{R}}_w(P)}{\Delta P} \leq \frac{\mathcal{R}_w(\tilde{\boldsymbol{\rho}}^* + \Delta\tilde{\boldsymbol{\rho}}^*, \boldsymbol{P}^* + \Delta\boldsymbol{P}^*) - \hat{\mathcal{R}}_w^{(\tilde{\boldsymbol{\rho}}^*+\Delta\tilde{\boldsymbol{\rho}}^*)}(P)}{\Delta P} \tag{B.3}$$

It is obvious when $\Delta P \to 0$, we have $\Delta\boldsymbol{P}^* \to \mathbf{0}$ and $\Delta\tilde{\boldsymbol{\rho}}^* \to \mathbf{0}$. Using the same approach as in the proof of Theorem 3.1, we have the equalities in (B.4) and (B.5). In (B.5), $\overset{(b)}{=}$ is obtained by removing the terms with $\tilde{\rho}_{k,n}^* + \Delta\tilde{\rho}_{k,n}^* = 0$, $\overset{(c)}{=}$ is obtained by removing the UEs whose rates are equal to their minimum rate requirements since they cannot lower power on their occupied subcarriers ($\mathcal{K}^*$ denotes the set of UEs whose rates are greater than their minimum rate requirements with the subcarrier and power allocation matrices $\tilde{\boldsymbol{\rho}}^* + \Delta\tilde{\boldsymbol{\rho}}^*$ and $\boldsymbol{P}^* + \Delta\boldsymbol{P}^*$), and $\overset{(d)}{=}$ follows the fact that $-\max_x f(x) = \min_x[-f(x)]$.

$$\begin{aligned} &\lim_{\Delta P\to 0} \frac{\hat{\mathcal{R}}_w^{(\tilde{\boldsymbol{\rho}}^*)}(P + \Delta P) - \mathcal{R}_w(\tilde{\boldsymbol{\rho}}^*, \boldsymbol{P}^*)}{\Delta P} \\ &= \max_{k\in\mathcal{K},n\in\mathcal{N}} \frac{\omega_k W\tilde{\rho}_{k,n}^* g_{k,n}\log_2 e}{\tilde{\rho}_{k,n}^* + g_{k,n}p_{k,n}^*}. \end{aligned} \tag{B.4}$$

$$
\begin{aligned}
&\lim_{\Delta P\to 0}\frac{\mathcal{R}_w(\tilde{\boldsymbol{\rho}}^*+\Delta\tilde{\boldsymbol{\rho}}^*,\boldsymbol{P}^*+\Delta\boldsymbol{P}^*)-\hat{\mathcal{R}}_w^{(\tilde{\rho}^*+\Delta\tilde{\rho}^*)}(P)}{\Delta P}\\
&=\lim_{\Delta P\to 0}\frac{\sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{N}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\log_2\left(1+\frac{g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n})}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}}\right)-\max\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{N}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\log_2\left(1+\frac{g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n}-\Delta p_{k,n})}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}}\right)}{\Delta P}\\
&\overset{(b)}{=}\lim_{\Delta P\to 0}\frac{\sum\limits_{\substack{k\in\mathcal{K},n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\log_2\left(1+\frac{g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n})}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}}\right)-\max\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{\substack{k\in\mathcal{K},n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\log_2\left(1+\frac{g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n}-\Delta p_{k,n})}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}}\right)}{\Delta P}\\
&\overset{(c)}{=}\lim_{\Delta P\to 0}\frac{\sum\limits_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\log_2\left(1+\frac{g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n})}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}}\right)-\max\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\log_2\left(1+\frac{g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n}-\Delta p_{k,n})}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}}\right)}{\Delta P}\\
&\overset{(d)}{=}\lim_{\Delta P\to 0}\frac{\min\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\log_2\left(\frac{(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})+g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n})}{(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})+g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n}-\Delta p_{k,n})}\right)}{\Delta P}\\
&=\lim_{\Delta P\to 0}\frac{\min\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\left[-\log_2\left(\frac{(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})+g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n}-\Delta p_{k,n})}{(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})+g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n})}\right)\right]}{\Delta P}\\
&=\lim_{\Delta P\to 0}\frac{\min\limits_{\Delta p_{k,n}\geq 0}\sum\limits_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})\left(\frac{g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}+g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n})}\Delta p_{k,n}+o(\Delta p_{k,n})\right)}{\Delta P}\\
&=\lim_{\Delta P\to 0}\min_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}>0}}\frac{\omega_k W(\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n})g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+\Delta\tilde{\rho}^*_{k,n}+g_{k,n}(p^*_{k,n}+\Delta p^*_{k,n})}\\
&=\min_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}>0}}\frac{\omega_k W\tilde{\rho}^*_{k,n}g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+g_{k,n}p^*_{k,n}}.
\end{aligned}
\tag{B.5}
$$

Notice that we naturally have

$$
\min_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}>0}}\frac{\omega_k W\tilde{\rho}^*_{k,n}g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+g_{k,n}p^*_{k,n}}\leq\max_{k\in\mathcal{K},n\in\mathcal{N}}\frac{\omega_k W\tilde{\rho}^*_{k,n}g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+g_{k,n}p^*_{k,n}}.
\tag{B.6}
$$

Combing (B.3), (B.4), (B.5), and (B.6), we have that

$$
\lim_{\Delta P\to 0}\frac{\hat{\mathcal{R}}_w(P+\Delta P)-\hat{\mathcal{R}}_w(P)}{\Delta P}=\max_{k\in\mathcal{K},n\in\mathcal{N}}\frac{\omega_k W\tilde{\rho}^*_{k,n}g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+g_{k,n}p^*_{k,n}}.
$$

Thus, $\frac{\mathrm{d}\hat{\mathcal{R}}_w(P)}{\mathrm{d}P}=\max_{k\in\mathcal{K},n\in\mathcal{N}}\frac{\omega_k W\tilde{\rho}^*_{k,n}g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+g_{k,n}p^*_{k,n}}$. Accordingly, Property (ii) and the remaining parts of Property (iii) can be proved in the same way as Theorem 3.1. Moreover, we have also

found that

$$\min_{\substack{k\in\mathcal{K}^*,n\in\mathcal{N}\\ s.t.\,\tilde{\rho}^*_{k,n}>0}} \frac{\omega_k W\tilde{\rho}^*_{k,n}g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+g_{k,n}p^*_{k,n}} = \max_{k\in\mathcal{K},n\in\mathcal{N}} \frac{\omega_k W\tilde{\rho}^*_{k,n}g_{k,n}\log_2 e}{\tilde{\rho}^*_{k,n}+g_{k,n}p^*_{k,n}}, \tag{B.7}$$

which indicates that for all the UEs whose rates are greater than their rate requirements, their occupied subcarriers have the same "water-level".

B.3 Proof of Theorem 3.3

Let $\boldsymbol{\rho}^{opt} \triangleq [\rho^{opt}_{k,n}]_{K\times N}$ and $\boldsymbol{P}^{opt} \triangleq [p^{opt}_{k,n}]_{K\times N}$ be the optimal subcarrier and power allocation matrices for (3.7), which indicates $\hat{\eta}^{DL}_{EE} = \frac{\sum_{k\in\mathcal{K}}\omega_k\sum_{n\in\mathcal{N}}\rho^{opt}_{k,n}r^{opt}_{k,n}}{\zeta\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}}p^{opt}_{k,n}+P_c}$. Obviously, for any fixed $\boldsymbol{\rho}\in\varrho$ and $\boldsymbol{P}\in\mathcal{P}$,

$$\begin{aligned}\eta^{DL}_{EE} &= \frac{\omega_1\sum\limits_{n\in\mathcal{N}}\rho_{1,n}r_{1,n}+\cdots+\omega_K\sum\limits_{n\in\mathcal{N}}\rho_{K,n}r_{K,n}}{\sum\limits_{n\in\mathcal{N}}\zeta p_{1,n}+\alpha_1P_c+\cdots+\sum\limits_{n\in\mathcal{N}}\zeta p_{K,n}+\alpha_KP_c}\\ &\geq \min_{k\in\mathcal{K}}\frac{\omega_k\sum\limits_{n\in\mathcal{N}}\rho_{k,n}r_{k,n}}{\sum\limits_{n\in\mathcal{N}}\zeta p_{k,n}+\alpha_kP_c},\end{aligned}$$

where $\boldsymbol{\alpha}\in\alpha$. And the equality holds if and only if

$$\frac{\omega_1\sum\limits_{n\in\mathcal{N}}\rho_{1,n}r_{1,n}}{\sum\limits_{n\in\mathcal{N}}\zeta p_{1,n}+\alpha_1P_c} = \cdots = \frac{\omega_K\sum\limits_{n\in\mathcal{N}}\rho_{K,n}r_{K,n}}{\sum\limits_{n\in\mathcal{N}}\zeta p_{K,n}+\alpha_KP_c},$$

which requires that $\alpha_k = \frac{\omega_k\sum_{n\in\mathcal{N}}\rho_{k,n}r_{k,n}}{\eta^{DL}_{EE}P_c} - \frac{\sum_{n\in\mathcal{N}}\zeta p_{k,n}}{P_c}$. Then it is obvious to have that

$$\begin{aligned}\hat{\eta}^{DL}_{EE} &= \max_{\boldsymbol{\rho}\in\varrho,\boldsymbol{P}\in\mathcal{P}}\eta^{DL}_{EE}\\ &\geq \max_{\boldsymbol{\rho}\in\varrho,\boldsymbol{P}\in\mathcal{P}}\left\{\min_{k\in\mathcal{K}}\frac{\omega_k\sum\limits_{n\in\mathcal{N}}\rho_{k,n}r_{k,n}}{\sum\limits_{n\in\mathcal{N}}\zeta p_{k,n}+\alpha_kP_c}\right\},\end{aligned}$$

where $\boldsymbol{\alpha}\in\alpha$. Thus, the RHS cannot exceed the LHS for any $\boldsymbol{\alpha}\in\alpha$. Furthermore, the equality always holds if the RHS uses the same $\boldsymbol{\rho}$ and $\boldsymbol{P}$ as $\boldsymbol{\rho}^{opt}$ and $\boldsymbol{P}^{opt}$ that achieve the LHS and sets $\{\alpha_k\}$ such that $\alpha_k = \alpha^{opt}_k \triangleq \frac{\omega_k\sum_{n\in\mathcal{N}}\rho^{opt}_{k,n}r^{opt}_{k,n}}{\hat{\eta}^{DL}_{EE}P_c} - \frac{\sum_{n\in\mathcal{N}}\zeta p^{opt}_{k,n}}{P_c}$. Clearly, straightforward addition with plugging in $\hat{\eta}^{DL}_{EE} = \frac{\sum_{k\in\mathcal{K}}\omega_k\sum_{n\in\mathcal{N}}\rho^{opt}_{k,n}r^{opt}_{k,n}}{\zeta\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}}p^{opt}_{k,n}+P_c}$ concludes that $\sum\limits_{k\in\mathcal{K}}\alpha^{opt}_k = 1$. This completes the proof.

APPENDIX C

PROOF FOR CHAPTER 4

C.1 Preliminaries on Effective Capacity

In this section, we first give the preliminaries on effective capacity for the statistical delay provisioning.

Statistical delay guarantees in networking with time-varying sources have been extensively studied by the theory of effective bandwidth [43]. Consider an arrival process $\{A(t), t \geq 0\}$ with $A(t)$ representing the amount of source data (in bits) over the time interval $[0, t)$. Assume that the Gärtner-Ellis limit of $A(t)$, which can be expressed as

$$\Lambda_{\mathrm{B}}(\vartheta) \triangleq \lim_{t\to\infty} \frac{1}{t} \ln \mathbb{E}\left(e^{\vartheta A(t)}\right), \tag{C.1}$$

exists for all $\vartheta \geq 0$. Then, the effective bandwidth of $A(t)$ is defined as

$$\mathrm{E}_{\mathrm{B}}(\vartheta) \triangleq \frac{\Lambda_{\mathrm{B}}(\vartheta)}{\vartheta} = \lim_{t\to\infty} \frac{1}{\vartheta t} \ln \mathbb{E}\left(e^{\vartheta A(t)}\right). \tag{C.2}$$

Using the theory of large deviations, for a queue of infinite buffer size served by a channel with constant service rate r, the steady-state delay violation probability for a given delay bound $D_{\max}$ satisfies

$$\Pr(D \geq D_{\max}) \approx p(r) e^{-\tilde{\vartheta}(r) D_{\max}}, \tag{C.3}$$

where $p(r) = \Pr(D > 0)$ is the probability that the buffer is nonempty and $\tilde{\vartheta}(r) = r\mathrm{E}_{\mathrm{B}}^{-1}(r)$ is the delay exponent for the delay violation probability. Therefore, the effective bandwidth concept provides the the minimum constant service rate $r_{\min}$ for a required delay exponent $\tilde{\vartheta}_{req}$ by solving the equation $\tilde{\vartheta}_{req} = r_{\min}\mathrm{E}_{\mathrm{B}}^{-1}(r_{\min})$.

Motivated by the theory of the effective bandwidth [43], effective capacity as a dual concept to the effective bandwidth has been proposed to characterize the transmission/service rate with statistical delay requirement in the scenario where the transmission

rates for reliable communication over wireless channels are time-varying [43, 44, 45]. Let $\{R[i], i = 1, 2, \cdots\}$ denote a discrete-time stationary and ergodic service process and $S[t] \triangleq \sum_{i=1}^{t} R[i]$ be the corresponding time-accumulated process, where $R[i]$ represents the amount of data served in the ith time frame. Assume that the Gärtner-Ellis limit of $S[t]$, which can be express as

$$\Lambda_{\mathrm{C}}(\theta) \triangleq \lim_{t\to\infty} \frac{1}{t} \ln \mathbb{E}\left(e^{\theta S[t]}\right), \tag{C.4}$$

exists for all $\theta \geq 0$. Then, the effective capacity of such a service process is given by [43]

$$\mathrm{E}_{\mathrm{C}}(\theta) \triangleq -\frac{\Lambda_{\mathrm{C}}(-\theta)}{\theta} = -\lim_{t\to\infty} \frac{1}{\theta t} \ln \mathbb{E}\left(e^{-\theta S[t]}\right). \tag{C.5}$$

When the service process, $\{R[i], i = 1, 2, \cdots\}$, is uncorrelated, the effective capacity simplifies to

$$\mathrm{E}_{\mathrm{C}}(\theta) = -\frac{1}{\theta} \ln \mathbb{E}\left(e^{-\theta R[i]}\right). \tag{C.6}$$

Furthermore, the effective capacity is a monotonically decreasing function of θ [84, 46].

Using the large deviation theory, for a queue of infinite buffer size and supplied by a source of constant arrival rate μ, the steady-state delay violation probability for a given delay bound $D_{\max}$ satisfies [43]

$$\Pr(D \geq D_{\max}) \approx p(\mu) e^{-\tilde{\theta}(\mu) D_{\max}}, \tag{C.7}$$

where $p(\mu) = \Pr(D > 0)$ is the probability that the buffer is nonempty and $\tilde{\theta}(\mu) = \mu \mathrm{E}_{\mathrm{C}}^{-1}(\mu)$ is the delay exponent for the delay violation probability. Hence, the effective capacity concept provides the the maximum constant arrival rate $\mu_{\max}$ for the service process to support with a required delay exponent $\tilde{\theta}_{req}$, which can be obtained by solving the equation $\tilde{\theta}_{req} = \mu_{\max} \mathrm{E}_{\mathrm{C}}^{-1}(\mu_{\max})$. Let $\theta_{req} \triangleq \mathrm{E}_{\mathrm{C}}^{-1}(\mu_{\max})$. The parameter θ (θ_{req}) instead of $\tilde{\theta}$ ($\tilde{\theta}_{req}$) approximately characterizes the steady-state delay violation probability. Clearly, a smaller θ indicates a looser delay constraint while a larger θ implies a more stringent delay guarantee. Particularly, when θ approaches zero, the arrival service has no specific delay requirement

and the effective capacity of the service process converges to the Shannon's ergodic capacity. On the other hand, as θ goes to infinity, the arrival service cannot tolerate any delay and the effective capacity becomes zero.

Therefore, the effective capacity of a service process is the maximum constant arrival rate that the service can support under the statistical delay constraint specified by the delay exponent θ. It can also be considered as the maximum throughput subject to such a delay requirement.

C.2 Proof of Lemma 4.1

Based on the fact that $\left(\int_a^b f(x,y)\mathrm{d}y\right)_x^{'} = \int_a^b f_x^{'}(x,y)\mathrm{d}y$, it is easy to prove that $\frac{\partial^2 \tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}},\boldsymbol{P}\right)}{\partial \tilde{\phi}_{k,n}^2} \leq 0$ and $\frac{\partial^2 \tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}},\boldsymbol{P}\right)}{\partial p_{k,n}^2} \leq 0$. Hence, $\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}},\boldsymbol{P}\right)$ is concave in $\tilde{\phi}_{k,n}$'s and $p_{k,n}$'s. Using the same technique in Lemma 1 in [70], we can prove that $\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}},\boldsymbol{P}\right)$ is further jointly concave in $\tilde{\phi}_{k,n}$'s and $p_{k,n}$'s. This completes the proof.

C.3 Proof of Lemma 4.2

Define

From Lemma 1, for $\tilde{\boldsymbol{\phi}}_1, \tilde{\boldsymbol{\phi}}_2, \boldsymbol{P}_1 \neq \boldsymbol{P}_2$, and $\lambda \in (0,1)$, we have that $\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\lambda\tilde{\boldsymbol{\phi}}_1+(1-\lambda)\tilde{\boldsymbol{\phi}}_2, \lambda\boldsymbol{P}_1+(1-\lambda)\boldsymbol{P}_2\right) \geq \lambda\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}}_1,\boldsymbol{P}_1\right)+(1-\lambda)\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}}_2,\boldsymbol{P}_2\right)$. On the other hand, for $\epsilon > 0$, there exist $\tilde{\boldsymbol{\phi}}_1^o$ and $\tilde{\boldsymbol{\phi}}_2^o$ such that $\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}}_i^o,\boldsymbol{P}_i\right) > \widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\boldsymbol{P}_i) - \epsilon$ for $i = 1,2$. Then, we have

$$\begin{aligned}
&\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\lambda\boldsymbol{P}_1+(1-\lambda)\boldsymbol{P}_2)\\
&= \max_{\tilde{\boldsymbol{\phi}}\in\tilde{\Phi}} \tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}}, \lambda\boldsymbol{P}_1+(1-\lambda)\boldsymbol{P}_2\right)\\
&\geq \tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\lambda\tilde{\boldsymbol{\phi}}_1^o + (1-\lambda)\tilde{\boldsymbol{\phi}}_2^o, \lambda\boldsymbol{P}_1+(1-\lambda)\boldsymbol{P}_2\right)\\
&\geq \lambda\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\lambda\tilde{\boldsymbol{\phi}}_1^o, \boldsymbol{P}_1\right)+(1-\lambda)\tilde{\mathrm{E}}_{\mathrm{C}}^{(\theta)}\left(\tilde{\boldsymbol{\phi}}_2^o, \boldsymbol{P}_2\right)\\
&> \lambda\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\boldsymbol{P}_1)+(1-\lambda)\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\boldsymbol{P}_2)-\epsilon.
\end{aligned} \tag{C.8}$$

Since (C.8) holds for any $\epsilon > 0$, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\lambda\boldsymbol{P}_1+(1-\lambda)\boldsymbol{P}_2) \geq \lambda\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\boldsymbol{P}_1)+(1-\lambda)\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\boldsymbol{P}_2)$. Hence, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\theta)}(\boldsymbol{P})$ is concave in $\boldsymbol{P}$.

Let $\boldsymbol{P}_j^*$ be the transmit power vector corresponding to $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_j)$, where $j = 1, 2, 3$. Without loss of generality, assume that $P_1 < P_2 < P_3$. Let $\boldsymbol{P}_2 = \frac{P_3-P_2}{P_3-P_1}\boldsymbol{P}_1^* + \frac{P_2-P_1}{P_3-P_1}\boldsymbol{P}_3^* = \delta\boldsymbol{P}_1^* + (1-\delta)\boldsymbol{P}_3^*$, where $\delta \triangleq \frac{P_3-P_2}{P_3-P_1} \in (0,1)$. Clearly, the sum of all components of $\boldsymbol{P}_2$ is equal to P_2. Then, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_2) = \widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{P}_2^*) \geq \widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{P}_2) \geq \delta\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{P}_1^*) + (1-\delta)\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(\boldsymbol{P}_3^*) = \delta\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_1) + (1-\delta)\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_3)$. Hence, $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$ is concave in P. This completes the proof of Lemma 2.

C.4 Proof of Theorem 4.1

Denote the superlevel set of $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$ as

$$\mathcal{S}_\alpha = \left\{ P \geq 0 \,|\, \widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P) \geq \alpha \right\}.$$

From [76], $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$ is quasiconcave in P if $\mathcal{S}_\alpha$ is convex for any real number α. When $\alpha < 0$, no points exist on the counter $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P) = \alpha$. When $\alpha \geq 0$, $\mathcal{S}_\alpha$ is equivalent to $\mathcal{S}_\alpha = \{P \geq 0 \,|\, \alpha\rho P + \alpha P_c - \tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P) \leq 0\}$. Since $\tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$ is concave in P, $\mathcal{S}_\alpha$ is convex in P. This completes the proof of Property (i).

On the other hand, with the concavity of $\tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$, it is obvious that $\lim\limits_{P\to\infty} \widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P) = \lim\limits_{P\to\infty} \frac{\tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\rho P + P_c} = \lim\limits_{P\to\infty} \frac{o(P)}{\rho P + P_c} = 0$. Thus, $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$ either decreases with P if $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}|_{P=P_0} \leq 0$ or first increases and then decreases with P if $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}|_{P=P_0} > 0$, where P_0 is the minimum total transmit power that can support all the QoS requirement. This completes the proof of Property (ii).

Since $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} = \lim\limits_{\Delta P\to 0} \frac{\frac{\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P+\Delta P)}{\rho(P+\Delta P)+P_c} - \frac{\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\rho P + P_c}}{T\Delta P} = \frac{\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} - \rho T\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{T(\rho P + P_c)}$, $\mathrm{sgn}\left(\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}\right) = \mathrm{sgn}\left(\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P} - \rho T\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)\right)$, where $\mathrm{sgn}(x)$ denotes the sign of x. And the derivative $\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)}{\mathrm{d}P}$ is derived in (C.9), where (a) is based on the fact that $\ln(1+x) \sim x$ when $x \to 0$ while (b) is based on the fact that $e^x - 1 \sim x$ and $\log_2(1+x) \sim x\log_2 e$ when $x \to 0$.

C.5 Proof of Property 4.2

According to [84, 46], $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(\boldsymbol{\phi}, \boldsymbol{P})$ is a strictly and monotonically decreasing function of all θ_k's. Without loss of generality, let $\boldsymbol{\theta}_1 \preceq \boldsymbol{\theta}_2$, where $\preceq$ is used as a component-wise

$$
\begin{aligned}
&\lim_{\Delta P\to 0}\frac{\left\{\max\limits_{\sum_{k=1}^{K}\Delta p_{k,n}=\Delta P}\sum\limits_{k=1}^{K}\sum\limits_{n=1}^{N}-\frac{1}{\theta_k}\ln\left[\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}+\Delta p_{k,n}}{\tilde{\phi}_{k,n}^{*}}\gamma_{k,n}\right)}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}\right]\right\}-\tilde{\mathrm{E}}_{\mathrm{C}}^{(\boldsymbol{\theta})}\left(\tilde{\boldsymbol{\phi}}^{*},\boldsymbol{P}^{*}\right)}{\Delta P}\\
&\overset{(a)}{=}\lim_{\Delta P\to 0}\frac{\max\limits_{\sum_{k=1}^{K}\Delta p_{k,n}=\Delta P}\sum\limits_{k=1}^{K}\sum\limits_{n=1}^{N}-\frac{1}{\theta_k}\frac{\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}+\Delta p_{k,n}}{\tilde{\phi}_{k,n}^{*}}\gamma_{k,n}\right)}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}-\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}}\right)}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}{\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}}\right)}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}}{\Delta P}\\
&\overset{(b)}{=}\lim_{\Delta P\to 0}\frac{\max\limits_{\sum_{k=1}^{K}\Delta p_{k,n}=\Delta P}\sum\limits_{k=1}^{K}\sum\limits_{n=1}^{N}\tilde{\phi}_{k,n}^{*}TB\log_2 e\frac{\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}}\right)}\frac{\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}+p_{k,n}^{*}\gamma_{k,n}}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}{\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}}\right)}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}\Delta p_{k,n}}{\Delta P}\\
&=\max_{k,n}\left\{\tilde{\phi}_{k,n}^{*}TB\log_2 e\frac{\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}}\right)}\frac{\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}+p_{k,n}^{*}\gamma_{k,n}}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}{\int_0^{\infty}e^{-\theta_k\tilde{\phi}_{k,n}^{*}TB\log_2\left(1+\frac{p_{k,n}^{*}\gamma_{k,n}}{\tilde{\phi}_{k,n}^{*}}\right)}f_{k,n}(\gamma_{k,n})\mathrm{d}\gamma_{k,n}}\right\}.
\end{aligned}
\tag{C.9}
$$

inequality. Then, we have that $\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_2)}(\boldsymbol{\phi},\boldsymbol{P})\le\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_1)}(\boldsymbol{\phi},\boldsymbol{P})$. Assume that $\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_2)}=\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_2)}(\boldsymbol{\phi}^{*},\boldsymbol{P}^{*})$. Then, we have that $\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_2)}(\boldsymbol{\phi}^{*},\boldsymbol{P}^{*})\le\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_1)}(\boldsymbol{\phi}^{*},\boldsymbol{P}^{*})\le\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_1)}$. Therefore, we have that $\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_2)}\le\widehat{\eta}_{EE}^{(\boldsymbol{\theta}_1)}$. This concludes the proof.

C.6 Proof of Property 4.3

Define $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,j})}(P)\triangleq\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}(P)$ and $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,j})}\triangleq\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta})}$, where the circuit power is $P_{c,j}\ge 0$ and $j=1,2$. Without loss of generality, let $P_{c,1}<P_{c,2}$. Let $P_{opt,j}$ be the total transmit power for achieving the maximum EE, i.e., $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,j})}=\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,j})}(P_{opt,j})$. Subsequently, we will prove $P_{opt,2}\ge P_{opt,1}$ based on (iii) in Theorem 1.

If $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,1})}(P_{opt,1})}{\mathrm{d}P}>0$, then we have $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,1})}(P_{opt,1})<\frac{1}{\rho T}\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_{opt,1})}{\mathrm{d}P}$. Thus, $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,2})}(P_{opt,1})<\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,1})}(P_{opt,1})<\frac{1}{\rho T}\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_{opt,1})}{\mathrm{d}P}$. Note that $\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P)$ is independent of $P_{c,j}$. Hence, $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,2})}(P_{opt,1})}{\mathrm{d}P}>0$ and thus $P_{opt,2}\ge P_{opt,1}$. Actually, in this case, $P_{opt,2}=P_{opt,1}=P_{\max}$.

If $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,1})}(P_{opt,1})}{\mathrm{d}P}=0$, then we have $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,1})}(P_{opt,1})=\frac{1}{\rho T}\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_{opt,1})}{\mathrm{d}P}$. Thus, $\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,2})}(P_{opt,1})<\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,1})}(P_{opt,1})=\frac{1}{\rho T}\frac{\mathrm{d}\widehat{\tilde{\mathrm{E}}}_{\mathrm{C}}^{(\boldsymbol{\theta})}(P_{opt,1})}{\mathrm{d}P}$. Hence, $\frac{\mathrm{d}\widehat{\tilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,2})}(P_{opt,1})}{\mathrm{d}P}>0$ and thus $P_{opt,2}\ge P_{opt,1}$.

If $\frac{\mathrm{d}\widehat{\widetilde{\eta}}_{EE}^{(\boldsymbol{\theta},P_{c,1})}(P_{opt,1})}{\mathrm{d}P} < 0$, then $P_{opt,1}$ is the minimum transmit power to ensure the delay requirement of all users. Hence, $P_{opt,2} \geq P_{opt,1}$.

To sum up, we have $P_{opt,2} \geq P_{opt,1}$. This completes the proof of Property 3.

APPENDIX D

PROOF FOR CHAPTER 5

D.1 Proof of Theorem 5.1

From [70], $\sum_{n\in\mathcal{N}} qW\varphi_{n,k}\log_2(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}})$ is jointly concave in $\boldsymbol{\varphi}_k$ and $\mathbf{p}_k$. Denote the superlevel sets of $\mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k)$ as

$$\mathcal{S}_\alpha = \{\boldsymbol{\theta}_k \succeq \mathbf{0}_N, \mathbf{p}_k \succeq \mathbf{0}_N \,|\, \mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k) \geq \alpha\}. \tag{D.1}$$

According to [76], $\mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k)$ is jointly quasiconcave in $\boldsymbol{\theta}_k$ and $\mathbf{p}_k$ if $\mathcal{S}_\alpha$ is jointly convex in $\boldsymbol{\theta}_k$ and $\mathbf{p}_k$ for any real number α. When $\alpha < 0$, no points exist on the counter $\mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k) = \alpha$. When $\alpha \geq 0$, $\mathcal{S}_\alpha$ is equivalent to $\mathcal{S}_\alpha = \{\boldsymbol{\theta}_k \succeq \mathbf{0}_N, \mathbf{p}_k \succeq \mathbf{0}_N \,|\, \alpha(p_c + \rho\sum_{n\in\mathcal{N}} p_{n,k}) - \sum_{n\in\mathcal{N}} qW\varphi_{n,k}\log_2(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}}) \leq 0\}$. Since $-\sum_{n\in\mathcal{N}} qW\varphi_{n,k}\log_2(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}}) + \alpha(p_c + \rho\sum_{n\in\mathcal{N}} p_{n,k})$ is jointly convex in $\boldsymbol{\theta}_k$ and $\mathbf{p}_k$, $\mathcal{S}_\alpha$ is jointly convex in $\boldsymbol{\theta}_k$ and $\mathbf{p}_k$. Hence, $\mathcal{E}_k(\boldsymbol{\theta}_k, \mathbf{p}_k)$ is jointly quasiconcave in $\boldsymbol{\theta}_k$ and $\mathbf{p}_k$

Since the unconstrained ξ_k depends only on $\boldsymbol{\varphi}_k$ and $\mathbf{p}_k$ but not other $\boldsymbol{\varphi}_{k'}$ and $\mathbf{p}_{k'}$ for $k' \neq k$, it is also jointly quasiconcave in $\boldsymbol{\Phi}$ and $\mathbf{P}$. Since the constraints in (5.3) and (5.8) consist of a convex set, $\xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k)$ under these convex constraints is still jointly quasiconcave in $\boldsymbol{\Phi}$ and $\mathbf{P}$. On the other hand, from [76], nonnegative weighted minimization of quasiconcave functions is still quasiconcave. Hence, the modified worst EE, $\underline{\xi}(\boldsymbol{\Phi}, \mathbf{P}) = \min_{k\in\mathcal{K}} w_k\xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k)$, under constraints in (5.3) and (5.8) is jointly quasiconcave in $\boldsymbol{\Phi}$ and $\mathbf{P}$. For any quasiconcave function, a local maximum is also a global maximum.

D.2 Proof of Theorem 5.2

As the unconstrained $\sum_{n\in\mathcal{N}} w_k qW\varphi_{n,k}\log_2(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}})$ is jointly concave in $\boldsymbol{\varphi}_k$ and $\mathbf{p}_k$, so is the unconstrained $J_k(\eta)$. Since the unconstrained $J_k(\eta)$ depends only on $\boldsymbol{\varphi}_k$ and $\mathbf{p}_k$ but not other $\boldsymbol{\varphi}_{k'}$ and $\mathbf{p}_{k'}$ for $k' \neq k$, it is also jointly concave in $\boldsymbol{\Phi}$ and $\mathbf{P}$. Thus, $J_k(\eta)$ under the convex constraints (5.3) and (5.8) is still strictly and jointly concave in $\boldsymbol{\Phi}$ and $\mathbf{P}$. On the other

hand, nonnegative weighted minimization of concave functions is still concave. Hence, $\underline{J}(\eta) \triangleq \min_{k\in\mathcal{K}} J_k(\eta)$ under constraints (5.3) and (5.8) is jointly concave in $\mathbf{\Phi}$ and $\mathbf{P}$. And it is obvious $J_k(\eta)$ strictly decreases with $\eta \geq 0$ and so do $\underline{J}(\eta)$ and $\widehat{\underline{J}}(\eta)$.

If $\eta = \underline{\xi}_o$, then $\underline{J}(\underline{\xi}_o) = \min_{k\in\mathcal{K}} \left\{(p_c + \rho \sum_{n\in\mathcal{N}} p_{n,k})\left(\frac{\sum_{n\in\mathcal{N}} w_k q W \varphi_{n,k} \log_2\left(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}}\right)}{p_c+\rho\sum_{n\in\mathcal{N}} p_{n,k}} - \underline{\xi}_o\right)\right\} \leq 0$, where the equality holds iff $\frac{\sum_{n\in\mathcal{S}_k^{\mathrm{v}}} w_k q W \log_2(1+p_{n,k}\gamma_{n,k})}{p_c+\rho\sum_{n\in\mathcal{S}_k^{\mathrm{v}}} p_{n,k}} = \underline{\xi}_o$ for at least one SU. Hence, $\widehat{\underline{J}}(\underline{\xi}_o) = 0$. Since $\widehat{\underline{J}}(\eta)$ strictly decreases with $\eta > 0$, $\widehat{\underline{J}}(\eta) > 0$ if $\eta < \underline{\xi}_o$ and $\widehat{\underline{J}}(\eta) < 0$ if $\eta > \underline{\xi}_o$. Thus, we have proved $\widehat{\underline{J}}(\eta) = 0$ iff $\eta = \underline{\xi}_o$.

D.3 Proof of Theorem 5.5

$\mathcal{D} = \mathcal{D}^{(0)}$ is obviously large enough to ensure that, for any $\mathbf{\Phi} \in \Phi$ and $\mathbf{P} \in \mathcal{P}$, there exists certain $(\boldsymbol{\alpha}'', \boldsymbol{\beta}'') \in \mathcal{D}$ such that the equalities in (5.19b) and (5.19c) strictly hold. Apparently, for $\mathcal{D} = \mathcal{D}^{(0)}$, the equalities in (5.19b) and (5.19c) all strictly hold when the optimality of problem (5.19) is achieved. Otherwise, we may increase β_k or decrease α_k for the inequalities to further improve the objective value. Thus, we have

$$H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^0)}, \mathbf{P}_o^{(\mathcal{D}^0)}, \boldsymbol{\alpha}_o^{(\mathcal{D}^0)}, \boldsymbol{\beta}_o^{(\mathcal{D}^0)}) = \overline{\underline{\xi}}(\mathbf{\Phi}_o^{(\mathcal{D}^0)}, \mathbf{P}_o^{(\mathcal{D}^0)}). \tag{D.2}$$

Moreover, $\exists\, \boldsymbol{\alpha}', \boldsymbol{\beta}'$ satisfying $(\boldsymbol{\alpha}', \boldsymbol{\beta}') \in \mathcal{D}^{(0)}$ such that

$$\overline{\underline{\xi}}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o) = H_{\mathcal{D}^0}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o, \boldsymbol{\alpha}', \boldsymbol{\beta}'). \tag{D.3}$$

On the other hand, we always have

$$\overline{\underline{\xi}}(\mathbf{\Phi}_o^{(\mathcal{D}^0)}, \mathbf{P}_o^{(\mathcal{D}^0)}) \leq \overline{\underline{\xi}}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o). \tag{D.4}$$

$$H_{\mathcal{D}^0}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o, \boldsymbol{\alpha}', \boldsymbol{\beta}') \leq H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^0)}, \mathbf{P}_o^{(\mathcal{D}^0)}, \boldsymbol{\alpha}_o^{(\mathcal{D}^0)}, \boldsymbol{\beta}_o^{(\mathcal{D}^0)}). \tag{D.5}$$

Then, from (D.2)-(D.5), we have

$$\begin{aligned} H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^0)}, \mathbf{P}_o^{(\mathcal{D}^0)}, \boldsymbol{\alpha}_o^{(\mathcal{D}^0)}, \boldsymbol{\beta}_o^{(\mathcal{D}^0)}) &= \overline{\underline{\xi}}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o) \\ &= \overline{\underline{\xi}}(\mathbf{\Phi}_o^{(\mathcal{D}^0)}, \mathbf{P}_o^{(\mathcal{D}^0)}) \\ &= H_{\mathcal{D}^0}(\overline{\mathbf{\Phi}}_o, \widetilde{\overline{\mathbf{P}}}_o, \boldsymbol{\alpha}', \boldsymbol{\beta}'). \end{aligned}$$

D.4 Concave Envelope of $h(\alpha,\beta) = \frac{\beta}{p_c+\rho\alpha}$

The concave envelope of $h(\alpha,\beta)$ follows the next theorem, which can be similarly proved as in [89, 86].

Theorem D.1 *The concave envelope of the function, $h(\alpha,\beta) = \frac{\beta}{p_c+\rho\alpha}$, on the compact convex set, $\mathcal{B} \triangleq \left\{(\alpha,\beta) \in \mathbb{R}^2 \,\middle|\, 0 \le \underline{\alpha} \le \alpha \le \overline{\alpha}, 0 \le \underline{\beta} \le \beta \le \overline{\beta}\right\}$, is*

$$h_{\mathcal{B}}^{\text{CE}}(\alpha,\beta) = \min\left(\frac{\beta}{p_c+\rho\underline{\alpha}} - \frac{\underline{\beta}(p_c+\rho\alpha)}{(p_c+\rho\underline{\alpha})(p_c+\rho\overline{\alpha})} + \frac{\underline{\beta}}{p_c+\rho\overline{\alpha}}, \frac{\beta}{p_c+\rho\overline{\alpha}} - \frac{\overline{\beta}(p_c+\rho\alpha)}{(p_c+\rho\underline{\alpha})(p_c+\rho\overline{\alpha})} + \frac{\overline{\beta}}{p_c+\rho\underline{\alpha}}\right). \tag{D.6}$$

Moreover, we have $h_{\mathcal{B}}^{\text{CE}}(\alpha,\beta) - h(\alpha,\beta) \to 0$, when $\overline{\alpha}-\underline{\alpha} \to 0$ and $\overline{\beta}-\underline{\beta} \to 0$ for any $(\alpha,\beta) \in \mathcal{B}$.

The fundamental definition of concave envelope ensures $h_{\mathcal{B}}^{\text{CE}}(\alpha,\beta) \ge h(\alpha,\beta)$ for every $(\alpha,\beta) \in \mathcal{B}$ and $h_{\mathcal{B}}^{\text{CE}}(\alpha,\beta)$ is jointly concave in α and β. Another straightforward observation about the Theorem above is when the rectangular region becomes small, the concave envelope, $h_{\mathcal{B}}^{\text{CE}}(\alpha,\beta)$, ensures to be tight for the original function, $h(\alpha,\beta)$, everywhere in the region. Apparently, the concave envelope of $h(\alpha_k,\beta_k) = \frac{\beta_k}{p_c+\rho\alpha_k}$ on the rectangular region, $\mathcal{B}_k^{(\mathcal{D})}$, is $h_{\mathcal{B}_k^{(\mathcal{D})}}^{\text{CE}}(\alpha_k,\beta_k)$.

D.5 Branching Strategy for the Proposed B&B Algorithm

For region $\mathcal{D}$, the longest edge of $\boldsymbol{\alpha}$ is dimension-ℓ as follows:

$$\ell = \arg\max_{k\in\mathcal{K}} \left(\overline{\alpha}_k^{(\mathcal{D})} - \underline{\alpha}_k^{(\mathcal{D})}\right).$$

Then, $\mathcal{B}_\ell^{(\mathcal{D})}$ is subdivided into $\dot{\mathcal{B}}_\ell^{(\mathcal{D})}$ and $\ddot{\mathcal{B}}_\ell^{(\mathcal{D})}$ by bisecting the range of α_ℓ as follows:

$$\dot{\mathcal{B}}_\ell^{(\mathcal{D})} \triangleq \left\{(\alpha_\ell,\beta_\ell) \,\middle|\, \underline{\alpha}_\ell^{(\mathcal{D})} \le \alpha_\ell < \frac{\underline{\alpha}_\ell^{(\mathcal{D})}+\overline{\alpha}_\ell^{(\mathcal{D})}}{2}, \underline{\beta}_\ell^{(\mathcal{D})} \le \beta_\ell \le \overline{\beta}_\ell^{(\mathcal{D})}\right\}.$$

$$\ddot{\mathcal{B}}_\ell^{(\mathcal{D})} \triangleq \left\{(\alpha_\ell,\beta_\ell) \,\middle|\, \frac{\underline{\alpha}_\ell^{(\mathcal{D})}+\overline{\alpha}_\ell^{(\mathcal{D})}}{2} \le \alpha_\ell \le \overline{\alpha}_\ell^{(\mathcal{D})}, \underline{\beta}_\ell^{(\mathcal{D})} \le \beta_\ell \le \overline{\beta}_\ell^{(\mathcal{D})}\right\}.$$

Accordingly, $\mathcal{D}$ is subdivided into $\dot{\mathcal{D}}$ and $\ddot{\mathcal{D}}$ by replacing $\mathcal{B}_\ell^{(\mathcal{D})}$ with $\dot{\mathcal{B}}_\ell^{(\mathcal{D})}$ and $\ddot{\mathcal{B}}_\ell^{(\mathcal{D})}$, respectively, as follows:

$$\dot{\mathcal{D}} \triangleq \mathcal{B}_1^{(\mathcal{D})} \times \cdots \mathcal{B}_{\ell-1}^{(\mathcal{D})} \times \dot{\mathcal{B}}_\ell^{(\mathcal{D})} \times \mathcal{B}_{\ell+1}^{(\mathcal{D})} \times \cdots \times \mathcal{B}_K^{(\mathcal{D})}.$$
$$\ddot{\mathcal{D}} \triangleq \mathcal{B}_1^{(\mathcal{D})} \times \cdots \mathcal{B}_{\ell-1}^{(\mathcal{D})} \times \ddot{\mathcal{B}}_\ell^{(\mathcal{D})} \times \mathcal{B}_{\ell+1}^{(\mathcal{D})} \times \cdots \times \mathcal{B}_K^{(\mathcal{D})}.$$

Apparently, $\dot{\mathcal{D}}$ and $\ddot{\mathcal{D}}$ consist of a non-overlapping partition of $\mathcal{D}$. In other words, we have $\dot{\mathcal{D}} \cup \ddot{\mathcal{D}} = \mathcal{D}$ and $\dot{\mathcal{D}} \cap \ddot{\mathcal{D}} = \varnothing$.

D.6 Proof of Theorem 5.6

From (D.6), it is easy to check, for a same (α, β), $h_{\mathcal{B}}^{\text{CE}}(\underline{\alpha}, \underline{\beta})$ is a decreasing function of $\underline{\alpha}$ and $\underline{\beta}$ while it is an increasing function of $\overline{\alpha}$ and $\overline{\beta}$. Hence, if shrinking a region, the concave envelope becomes tighter/smaller. Thus, $\max_j \widehat{H}_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})})$ is a decreasing function of i, since the branching process shrinks $\mathcal{D}^{(i,j)}$'s for all the survival branches.

On the other hand, the branching process based on the branching strategy in Appendix D.5 makes the largest edge of $\boldsymbol{\alpha}$ of each survival branch decrease in the order of $(\frac{1}{2})^i$. As the iteration goes on ($i \to \infty$), all the survival branches converge to some subregions with their respective fixed $\boldsymbol{\alpha}$'s (in fact, $\boldsymbol{\alpha}_o^{(\mathcal{D})}$'s), making $\widehat{H}_{\mathcal{D}^{(i,j)}}(\mathbf{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \to H_{\mathcal{D}^{(i,j)}}(\mathbf{\Phi}, \mathbf{P}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ for any $\boldsymbol{\beta}$ in the survival subregions. Note that all the regions/branches containing any optimal solution to (5.19) will never be pruned by the pruning process since their upper bounds are certainly no less than the maximum of (5.19). Hence, we have $\lim_{i\to\infty} \max_j \widehat{H}_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})}) = \lim_{i\to\infty} \max_j H_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})}) = H_{\mathcal{D}^0}(\mathbf{\Phi}_o^{(\mathcal{D}^0)}, \mathbf{P}_o^{(\mathcal{D}^0)}, \boldsymbol{\alpha}_o^{(\mathcal{D}^0)}, \boldsymbol{\beta}_o^{(\mathcal{D}^0)})$. In practice, a reasonable stopping criterion is

$$\max_j \widehat{H}_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})}) - \max_j H_{\mathcal{D}^{(i,j)}}(\widehat{\mathbf{\Phi}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\mathbf{P}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\alpha}}_o^{(\mathcal{D}^{(i,j)})}, \widehat{\boldsymbol{\beta}}_o^{(\mathcal{D}^{(i,j)})}) < \epsilon.$$

D.7 Proof of Theorem 5.7

Without loss of generality, here we assume $\boldsymbol{\varphi}_k \succ \mathbf{0}_N^T$. That is, $\varphi_{n,k} > 0$ for $\forall\, n \in \mathcal{N}$. This is because, for $\varphi_{n,k} = 0$, we can artificially let $\varphi_{n,k} = \tau$, where τ is a relatively small positive number or even $\tau \to 0^+$, to make $\varphi_{n,k} > 0$. All the involved optimization results practically remain the same mainly because $\lim_{\tau\to 0^+} \tau W \log_2\left(1 + \frac{p_{n,k}\gamma_{n,k}}{\tau}\right) = 0$.

As proved in Theorem 1 [85], for any given $\boldsymbol{\varphi}_k$, $\xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k)$ is strictly quasiconcave in the total transmit power $p_k (= \mathbf{1}_N^T \cdot \mathbf{p}_k)$. Let $\widehat{\boldsymbol{p}}_k = [\widehat{p}_{1,k}, \cdots, \widehat{p}_{N,k}]^T$ be the (optimal) transmit power vector that maximizes $\xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k)$ among all the vectors, $\mathbf{p}_k \in \mathcal{P}_k$, which have a total power of p_k. That is,

$$\widehat{\boldsymbol{p}}_k \equiv \arg \underbrace{\max_{\substack{\mathbf{1}_N^T \cdot \mathbf{p}_k = p_k, \\ \mathbf{p}_k \in \mathcal{P}_k}} \xi_k(\boldsymbol{\varphi}_k, \mathbf{p}_k)}_{\triangleq\, \xi_k^{(\boldsymbol{\varphi}_k)}(p_k)}. \tag{D.7}$$

Then, like Property (iii) in Theorem 1 and Property (iii) in Theorem 2 [85], we can similarly prove

$$\frac{\mathrm{d}\xi_k^{(\boldsymbol{\varphi}_k)}(p_k)}{\mathrm{d}p_k} \begin{cases} > 0, & \text{if } \xi_k^{(\boldsymbol{\varphi}_k)}(p_k) < \max\limits_{n\in\mathcal{N}} \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}\widehat{p}_{n,k}}{\varphi_{n,k}}\right)} \\ = 0, & \text{if } \xi_k^{(\boldsymbol{\varphi}_k)}(p_k) = \max\limits_{n\in\mathcal{N}} \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}\widehat{p}_{n,k}}{\varphi_{n,k}}\right)} \\ < 0, & \text{if } \xi_k^{(\boldsymbol{\varphi}_k)}(p_k) > \max\limits_{n\in\mathcal{N}} \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}\widehat{p}_{n,k}}{\varphi_{n,k}}\right)}. \end{cases} \tag{D.8}$$

Moreover, as a result of the implicit water-filling process when obtaining $\widehat{\boldsymbol{p}}_k$ in (D.7), we even have, for any $n' \in \mathcal{N}_+ \triangleq \{n \in \mathcal{N} | \varphi_{n,k} > 0, \widehat{p}_{n,k} > 0\}$,

$$\max_{n\in\mathcal{N}} \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}\widehat{p}_{n,k}}{\varphi_{n,k}}\right)} \equiv \frac{qW\gamma_{n',k}\log_2 e}{\rho\left(1+\frac{\gamma_{n',k}\widehat{p}_{n',k}}{\varphi_{n',k}}\right)}. \tag{D.9}$$

Note that we have assumed that $\boldsymbol{\varphi}_k \succ \mathbf{0}_N^T$ at the beginning.

Based on (D.8), (D.9), and the definition of $\mathbf{p}_k(\boldsymbol{\varphi}_k)$, we then have $\frac{\mathrm{d}\xi_k^{(\boldsymbol{\varphi}_k)}(p_k(\boldsymbol{\varphi}_k))}{\mathrm{d}p_k} \geq 0$ and

$$(\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k(\boldsymbol{\varphi}_k)))' = \underbrace{\frac{\sum_{n\in\mathcal{N}} qW\varphi_{n,k}\log_2\left(1+\frac{p_{n,k}(\boldsymbol{\varphi}_k)\gamma_{n,k}}{\varphi_{n,k}}\right)}{p_c+\rho\sum_{n\in\mathcal{N}} p_{n,k}(\boldsymbol{\varphi}_k)}}_{=\,\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k(\boldsymbol{\varphi}_k))}\cdot\frac{1}{\varphi_{n,k}} - \frac{\sum_{n\in\mathcal{N}} qWp_{n,k}(\boldsymbol{\varphi}_k)\gamma_{n,k}\log_2 e}{\left(p_c+\rho\sum_{n\in\mathcal{N}} p_{n,k}(\boldsymbol{\varphi}_k)\right)\left(p_{n,k}(\boldsymbol{\varphi}_k)\gamma_{n,k}+\varphi_{k,n}\right)} + o(\varphi_{n,k}). \tag{D.12}$$

$$\left(\frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}p_{n,k}(\boldsymbol{\varphi}_k)}{\varphi_{n,k}}\right)}\right)' = \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}p_{n,k}(\boldsymbol{\varphi}_k)}{\varphi_{n,k}}\right)}\cdot\frac{1}{\varphi_{k,n}} + o(\varphi_{n,k}). \tag{D.13}$$

---

thus

$$\begin{aligned}\overbrace{\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k(\boldsymbol{\varphi}_k))}^{\equiv\,\xi_k^{(\boldsymbol{\varphi}_k)}(p_k(\boldsymbol{\varphi}_k))} &= \frac{\sum_{n\in\mathcal{N}} qW\varphi_{n,k}\log_2\left(1+\frac{p_{n,k}(\boldsymbol{\varphi}_k)\gamma_{n,k}}{\varphi_{n,k}}\right)}{p_c+\rho\sum_{n\in\mathcal{N}} p_{n,k}(\boldsymbol{\varphi}_k)}\\ &\begin{cases} = \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}p_{n,k}(\boldsymbol{\varphi}_k)}{\varphi_{n,k}}\right)}, \forall n\in\mathcal{N}_+, & \text{if } p_k(\boldsymbol{\varphi}_k) < p_k^{\max}\\ \le \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}p_{n,k}(\boldsymbol{\varphi}_k)}{\varphi_{n,k}}\right)}, \forall n\in\mathcal{N}_+, & \text{if } p_k(\boldsymbol{\varphi}_k) = p_k^{\max}. \end{cases}\end{aligned} \tag{D.10}$$

Anyway, we always have

$$\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k(\boldsymbol{\varphi}_k)) \le \frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}p_{n,k}(\boldsymbol{\varphi}_k)}{\varphi_{n,k}}\right)}, \forall n\in\mathcal{N}_+. \tag{D.11}$$

Next, we calculate the derivative of $\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k(\boldsymbol{\varphi}_k))$ and $\frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}p_{n,k}(\boldsymbol{\varphi}_k)}{\varphi_{n,k}}\right)}$ with respect to $\varphi_{n,k}$ in a relatively small neighborhood around $\varphi_{n,k}$ for $n\in\mathcal{N}_+$ in (D.12) and (D.13), respectively, where $o(\varphi_{n,k})$ denotes a higher-order infinitesimal of $\varphi_{n,k}$.

According to (D.11)-(D.13), we have

$$(\xi_k(\boldsymbol{\varphi}_k,\mathbf{p}_k(\boldsymbol{\varphi}_k)))' \le \left(\frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}\widehat{p}_{n,k}}{\varphi_{n,k}}\right)}\right)',\ \forall\, n\in\mathcal{N}, \tag{D.14}$$

Note that both the two sides in (D.14) are zeros for $n\in\mathcal{N}_0 \triangleq \{n\in\mathcal{N}|\varphi_{n,k}>0,\widehat{p}_{n,k}=0\}$. This indicates $\frac{qW\gamma_{n,k}\log_2 e}{\rho\left(1+\frac{\gamma_{n,k}\widehat{p}_{n,k}}{\varphi_{n,k}}\right)}$ increases faster than $\xi_k^{(\boldsymbol{\varphi}_k)}(p_k(\boldsymbol{\varphi}_k))$ with respect to $\varphi_{n,k}$. Hence, if any $\varphi_{n,k}$ increases, one of the first two cases in (D.8) will hold with the previous $p_k(\boldsymbol{\varphi}_k)$,

indicating increasing the total transmit power will lead to a higher EE. Hence, for $\dot{\boldsymbol{\varphi}}_k \preceq \ddot{\boldsymbol{\varphi}}_k$, we have $p_k(\dot{\boldsymbol{\varphi}}_k) \leq p_k(\ddot{\boldsymbol{\varphi}}_k)$.

It is easy to prove $\varphi_{n,k}\log_2\left(1+\frac{p_{n,k}\gamma_{n,k}}{\varphi_{n,k}}\right)$ is an increasing function of $\varphi_{n,k}$. Hence, we have

$$\begin{aligned}
\widetilde{r}_k(\dot{\boldsymbol{\varphi}}_k) &= \sum_{n\in\mathcal{N}} qW\dot{\varphi}_{n,k}\log_2(1+\frac{p_{n,k}(\dot{\boldsymbol{\varphi}}_k)\gamma_{n,k}}{\dot{\varphi}_{n,k}}) \\
&\leq \sum_{n\in\mathcal{N}} qW\ddot{\varphi}_{n,k}\log_2(1+\frac{p_{n,k}(\dot{\boldsymbol{\varphi}}_k)\gamma_{n,k}}{\ddot{\varphi}_{n,k}}) \\
&\leq \sum_{n\in\mathcal{N}} qW\ddot{\varphi}_{n,k}\log_2(1+\frac{p_{n,k}(\ddot{\boldsymbol{\varphi}}_k)\gamma_{n,k}}{\ddot{\varphi}_{n,k}}) \\
&= \widetilde{r}_k(\ddot{\boldsymbol{\varphi}}_k).
\end{aligned}$$

D.8 Proof of Theorem 5.8

Clearly, since $\mathbf{p}_k(\boldsymbol{\varphi}_k) \in \mathcal{P}_k$ for any indicator vector $\boldsymbol{\varphi}_k \in \Phi_k$, we have $p_k(\mathbf{1}_N) \leq p_k^{\max}$ and $\widetilde{r}_k(\mathbf{1}_N) \leq \widetilde{r}_k^{\max}$. Hence, $\widetilde{\mathcal{B}}_k^{(0)} \subseteq \mathcal{D}_k^{(0)}$ and $\widetilde{\mathcal{D}}^{(0)} \subseteq \mathcal{D}^{(0)}$.

Let $[\boldsymbol{\varphi}_{o,1}^{(\mathcal{D}^{(0)})},\cdots,\boldsymbol{\varphi}_{o,K}^{(\mathcal{D}^{(0)})}] = \boldsymbol{\Phi}_o^{(\mathcal{D}^{(0)})}$, $[\mathbf{p}_{o,1}^{(\mathcal{D}^{(0)})},\cdots,\mathbf{p}_{o,K}^{(\mathcal{D}^{(0)})}] = \mathbf{P}_o^{(\mathcal{D}^{(0)})}$, $[\alpha_{o,1}^{(\mathcal{D}^{(0)})},\cdots,\alpha_{o,K}^{(\mathcal{D}^{(0)})}]^T = \boldsymbol{\alpha}_o^{(\mathcal{D}^{(0)})}$, and $[\beta_{o,1}^{(\mathcal{D}^{(0)})},\cdots,\beta_{o,K}^{(\mathcal{D}^{(0)})}]^T = \boldsymbol{\beta}_o^{(\mathcal{D}^{(0)})}$. Because of the equivalence of (5.12) and (5.19) when $\mathcal{D} = \mathcal{D}^{(0)}$, we certainly have $\mathbf{p}_{o,k}^{(\mathcal{D}^{(0)})} = \mathbf{p}_k(\boldsymbol{\varphi}_{o,k}^{(\mathcal{D}^{(0)})})$. Since constraints (5.19b) and (5.19c) must hold when the optimality of (5.19) is achieved, we have $\alpha_{o,k}^{(\mathcal{D}^{(0)})} = \mathbf{1}_N^T \cdot \mathbf{p}_{o,k}^{(\mathcal{D}^{(0)})} = \mathbf{1}_N^T \cdot \mathbf{p}_k(\boldsymbol{\varphi}_{o,k}^{(\mathcal{D}^{(0)})}) = p_k(\boldsymbol{\varphi}_{o,k}^{(\mathcal{D}^{(0)})}) \leq p_k(\mathbf{1}_N)$ and $\beta_{o,k}^{(\mathcal{D}^{(0)})} = \widetilde{r}_k(\boldsymbol{\varphi}_{o,k}^{(\mathcal{D}^{(0)})}) \leq \widetilde{r}(\mathbf{1}_N)$. Hence, we have $(\boldsymbol{\alpha}_o^{(\mathcal{D}^{(0)})}, \boldsymbol{\beta}_o^{(\mathcal{D}^{(0)})}) \in \widetilde{\mathcal{D}}^{(0)}$.

APPENDIX E

PROOF FOR CHAPTER 6

E.1 Proof of Theorem 6.1

Let $\varpi(\overline{\mathbf{p}}_n) \triangleq \frac{T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n)}{4p_c^D|\mathcal{H}_n|+4p_c^S+p\sum_{i\in\mathcal{H}_n}(p_{n,i}^{(1)}+p_{n,i}^{(2)})}$. Apparently, we have $\varpi(\overline{\mathbf{p}}_n) = \frac{\sum_{i\in\mathcal{H}_n}\left(\log_2\left(1+\gamma_{n,i}^{(1)}\right)+\log_2\left(1+\gamma_{n,i}^{(2)}\right)\right)}{4p_c^D|\mathcal{H}_n|+4p_c^S+p\sum_{i\in\mathcal{H}_n}(p_{n,i}^{(1)}+p_{n,i}^{(2)})} - \xi_n$. Under constraints in (6.5b), we obviously have

$$\begin{cases} \max_{\overline{\mathbf{p}}_n} \varpi(\overline{\mathbf{p}}_n) < 0 & \text{if} \quad \xi_n > \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)} \\ \max_{\overline{\mathbf{p}}_n} \varpi(\overline{\mathbf{p}}_n) = 0 & \text{if} \quad \xi_n = \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)} \\ \max_{\overline{\mathbf{p}}_n} \varpi(\overline{\mathbf{p}}_n) > 0 & \text{if} \quad \xi_n < \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}. \end{cases}$$

Moreover, we further have

$$\begin{cases} T_n^{(\mathcal{H}_n)}(\xi_n) = \max_{\overline{\mathbf{p}}_n} T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n) < 0 & \text{if} \quad \max_{\overline{\mathbf{p}}_n} \varpi(\overline{\mathbf{p}}_n) < 0 \\ T_n^{(\mathcal{H}_n)}(\xi_n) = \max_{\overline{\mathbf{p}}_n} T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n) = 0 & \text{if} \quad \max_{\overline{\mathbf{p}}_n} \varpi(\overline{\mathbf{p}}_n) = 0 \\ T_n^{(\mathcal{H}_n)}(\xi_n) = \max_{\overline{\mathbf{p}}_n} T_n^{(\mathcal{H}_n,\xi_n)}(\overline{\mathbf{p}}_n) > 0 & \text{if} \quad \max_{\overline{\mathbf{p}}_n} \varpi(\overline{\mathbf{p}}_n) > 0. \end{cases}$$

The completes the proof of Theorem 6.1.

E.2 Proof of Theorem 6.2

For brevity, we simply the notations of $\gamma_{n,i}^{(1)}$ and $\gamma_{n,i}^{(2)}$ as

$$\gamma_{n,i}^{(1)} = \frac{\vartheta p_{n,i}^{(2)}}{\alpha + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2} \text{ and } \gamma_{n,i}^{(2)} = \frac{\vartheta p_{n,i}^{(1)}}{\beta + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2},$$

respectively, where $\vartheta \triangleq q_i|h_{n,i}^{(1)}|^2|h_{n,i}^{(2)}|^2$, $\alpha \triangleq 1 + q_i|h_{n,i}^{(1)}|^2$, and $\beta \triangleq 1 + q_i|h_{n,i}^{(2)}|^2$.

First, we prove the monotonicity of $C_{n,i}(p_{n,i})$. Without loss of generality, let $0 \le p_{n,i} < \overline{p}_{n,i}$. Then, if $p_{n,i} = 0$, we obviously have $0 = C_{n,i}(p_{n,i}) < C_{n,i}(\overline{p}_{n,i})$; otherwise,

$$C_{n,i}(p_{n,i}) = R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i})) < R_{n,i}(\frac{\overline{p}_{n,i}}{p_{n,i}}\mathbf{p}_{n,i}^*(p_{n,i})) \le C_{n,i}(\overline{p}_{n,i}),$$

where the strict inequality is from straightforward algebra manipulation while the last inequality is based on the simple fact that $\mathbf{1}_2^T \cdot \frac{\overline{p}_{n,i}}{p_{n,i}}\mathbf{p}_{n,i}^*(p_{n,i}) = \overline{p}_{n,i}$. Therefore, $C_{n,i}(p_{n,i}) < C_{n,i}(\overline{p}_{n,i})$ for $0 \le p_{n,i} < \overline{p}_{n,i}$.

Next, we prove the concavity of $C_{n,i}(p_{n,i})$. The first-order partial derivative of $R_{n,i}(\mathbf{p}_{n,i})$ w.r.t. $p_{n,i}^{(j)}$ is

$$\frac{\partial R_{n,i}(\mathbf{p}_{n,i})}{\partial p_{n,i}^{(j)}} = \log_2 e\Big[\frac{(j-1)\vartheta + |h_{n,i}^{(j)}|^2}{\alpha + \vartheta p_{n,i}^{(2)} + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2} - \frac{|h_{n,i}^{(j)}|^2}{\alpha + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2} + \frac{(2-j)\vartheta + |h_{n,i}^{(j)}|^2}{\beta + \vartheta p_{n,i}^{(1)} + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2} - \frac{|h_{n,i}^{(j)}|^2}{\beta + p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2 + p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2}\Big].$$

Moreover, the second-order partial derivatives are

$$\frac{\partial^2 R_{n,i}(\mathbf{p}_{n,i})}{\partial (p_{n,i}^{(j)})^2} = \log_2 e\Big[- \frac{\left((j-1)\vartheta + |h_{n,i}^{(j)}|^2\right)^2}{(\alpha+\vartheta p_{n,i}^{(2)}+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} + \frac{|h_{n,i}^{(j)}|^4}{(\alpha+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} + \frac{|h_{n,i}^{(j)}|^4}{(\beta+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} - \frac{\left((2-j)\vartheta+|h_{n,i}^{(j)}|^2\right)^2}{(\beta+\vartheta p_{n,i}^{(1)}+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2}\Big],$$

$$\frac{\partial^2 R_{n,i}(\mathbf{p}_{n,i})}{\partial p_{n,i}^{(1)} \partial p_{n,i}^{(2)}} = \log_2 e\Big[- \frac{|h_{n,i}^{(1)}|^2(\vartheta + |h_{n,i}^{(2)}|^2)}{(\alpha+\vartheta p_{n,i}^{(2)}+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} + \frac{|h_{n,i}^{(1)}|^2|h_{n,i}^{(2)}|^2}{(\alpha+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} + \frac{|h_{n,i}^{(1)}|^2|h_{n,i}^{(2)}|^2}{(\beta+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} - \frac{(\vartheta+|h_{n,i}^{(1)}|^2)|h_{n,i}^{(2)}|^2}{(\beta+\vartheta p_{n,i}^{(1)}+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2}\Big].$$

Accordingly, we have

$$\begin{aligned} & 2p_{n,i}^{(1)}p_{n,i}^{(2)}\frac{\partial^2 R_{n,i}(\mathbf{p}_{n,i})}{\partial p_{n,i}^{(1)}\partial p_{n,i}^{(2)}} + \sum_{j\in\mathcal{J}}\left(p_{n,i}^{(j)}\right)^2\frac{\partial^2 R_{n,i}(\mathbf{p}_{n,i})}{\partial\left(p_{n,i}^{(j)}\right)^2} \\ = \; & \log_2 e\Big[- \frac{(\vartheta p_{n,i}^{(2)} + |h_{n,i}^{(1)}|^2 p_{n,i}^{(1)} + |h_{n,i}^{(2)}|^2 p_{n,i}^{(2)})^2}{(\alpha+\vartheta p_{n,i}^{(2)}+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} + \frac{(|h_{n,i}^{(1)}|^2 p_{n,i}^{(1)} + |h_{n,i}^{(2)}|^2 p_{n,i}^{(2)})^2}{(\alpha+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} \\ & + \frac{(|h_{n,i}^{(1)}|^2 p_{n,i}^{(1)} + |h_{n,i}^{(2)}|^2 p_{n,i}^{(2)})^2}{(\beta+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2} - \frac{(\vartheta p_{n,i}^{(1)}+|h_{n,i}^{(2)}|^2 p_{n,i}^{(2)} + |h_{n,i}^{(1)}|^2 p_{n,i}^{(1)})^2}{(\beta+\vartheta p_{n,i}^{(1)}+p_{n,i}^{(1)}|h_{n,i}^{(1)}|^2+p_{n,i}^{(2)}|h_{n,i}^{(2)}|^2)^2}\Big] < 0, \end{aligned} \tag{E.1}$$

where the last inequality is due to the simple fact that $\frac{x^2}{(\alpha+x)^2}$ and $\frac{x^2}{(\beta+x)^2}$ are both increasing functions of $x > 0$.

According to the continuity and the strict monotonicity of $C_{n,i}(p_{n,i})$, it can be inferred that at least one of $\frac{\partial R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial p_{n,i}^{(1)}}$ and $\frac{\partial R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial p_{n,i}^{(2)}}$ is strictly greater than zero. Moreover, the

relative size of $\frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}}$ and $\frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i}}$ can be categorized by three possible cases:

$$\begin{cases} (i) & \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}} = \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i}} > 0 \\ (ii) & \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}} > \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i}} \ \& \ \mathbf{p}^*_{n,i}(p_{n,i}) = (p_{n,i}, 0)^T \\ (iii) & \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}} < \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i}} \ \& \ \mathbf{p}^*_{n,i}(p_{n,i}) = (0, p_{n,i})^T. \end{cases}$$

Otherwise, if $\frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}} > \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i}}$ while $p^{(2),*}_{n,i} > 0$, the value of $R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))$ can be improved by decreasing $p^{(2),*}_{n,i}$ and increasing $p^{(1),*}_{n,i}$, considering $\frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}}$ ensures to be strictly greater than zero when $\frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}} > \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i}}$. This contradicts the assumption that $\mathbf{p}^*_{n,i}(p_{n,i})$ is the optimal power vector.

For Case (i), by continuing taking partial derivatives, we can further have

$$\frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial \left(p^{(1)}_{n,i}\right)^2} = \frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i} \partial p^{(1)}_{n,i}} = \frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i} \partial p^{(2)}_{n,i}} = \frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial \left(p^{(2)}_{n,i}\right)^2}. \tag{E.2}$$

Hence, together with (E.1), we have

$$(p^{(1),*}_{n,i}(p_{n,i}) + p^{(2),*}_{n,i}(p_{n,i}))^2 \frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial \left(p^{(1)}_{n,i}\right)^2} < 0. \tag{E.3}$$

Meanwhile, the derivative of $C_{n,i}(p_{n,i})$ w.r.t $p_{n,i}$ satisfies

$$\begin{aligned} \frac{\mathrm{d}C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}} &= \frac{\partial C_{n,i}(p_{n,i})}{\partial p^{(1),*}_{n,i}(p_{n,i})} \frac{\partial p^{(1),*}_{n,i}(p_{n,i})}{\partial p_{n,i}} + \frac{\partial C_{n,i}(p_{n,i})}{\partial p^{(2),*}_{n,i}(p_{n,i})} \frac{\partial p^{(2),*}_{n,i}(p_{n,i})}{\partial p_{n,i}} \\ &= \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}} \frac{\partial p^{(1),*}_{n,i}(p_{n,i})}{\partial p_{n,i}} + \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(2)}_{n,i}} \frac{\partial p^{(2),*}_{n,i}(p_{n,i})}{\partial p_{n,i}} \\ &= \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}} \frac{\partial (p^{(1),*}_{n,i}(p_{n,i}) + p^{(2),*}_{n,i}(p_{n,i}))}{\partial p_{n,i}} = \frac{\partial R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i}}, \end{aligned} \tag{E.4}$$

where the first equality is based on the chain rule for composite functions while the last equality is because of $p^{(1),*}_{n,i}(p_{n,i}) + p^{(2),*}_{n,i}(p_{n,i}) = p_{n,i}$. Similarly, the second-order derivative of $C_{n,i}(p_{n,i})$ w.r.t $p_{n,i}$ satisfies

$$\frac{\mathrm{d}^2 C_{n,i}(p_{n,i})}{\mathrm{d}p^2_{n,i}} = \frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^2_{\dot{n},i}} \frac{\partial p^{(1),*}_{n,i}(p_{n,i})}{\partial p_{n,i}} + \frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial p^{(1)}_{n,i} \partial p^{(2)}_{n,i}} \frac{\partial p^{(2),*}_{n,i}(p_{n,i})}{\partial p_{n,i}} = \frac{\partial^2 R_{n,i}(\mathbf{p}^*_{n,i}(p_{n,i}))}{\partial \left(p^{(1)}_{n,i}\right)^2}. \tag{E.5}$$

Using (E.3) and (E.5), we have $\frac{\mathrm{d}^2 C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}^2} < 0$ for $p_{n,i} > 0$.

For Case (ii), we simply have

$$\frac{\mathrm{d}C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}} = \frac{\partial C_{n,i}(p_{n,i})}{\partial p_{n,i}^{(1),*}(p_{n,i})} = \frac{\partial R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial p_{n,i}^{(1)}}, \tag{E.6}$$

$$\frac{\mathrm{d}^2 C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}^2} = \left.\frac{\partial^2 R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial \left(p_{n,i}^{(1)}\right)^2}\right|_{\mathbf{p}_{n,i}^*(p_{n,i})=(p_{n,i},0)^T} < 0, \tag{E.7}$$

where the strict inequality is due to (E.1).

Similarly, for Case (iii), we have

$$\frac{\mathrm{d}C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}} = \frac{\partial C_{n,i}(p_{n,i})}{\partial p_{n,i}^{(2),*}(p_{n,i})} = \frac{\partial R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial p_{n,i}^{(2)}}, \tag{E.8}$$

$$\frac{\mathrm{d}^2 C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}^2} = \left.\frac{\partial^2 R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial \left(p_{n,i}^{(2)}\right)^2}\right|_{\mathbf{p}_{n,i}^*(p_{n,i})=(0,p_{n,i})^T} < 0. \tag{E.9}$$

Therefore, the strict concavity of $C_{n,i}(p_{n,i})$ is sustained for all the three possible cases. From (E.4), (E.6), and (E.8), we have $\frac{\mathrm{d}C_{n,i}(p_{n,i})}{\mathrm{d}p_{n,i}} = \max\{\frac{\partial R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial p_{n,i}^{(1)}}, \frac{\partial R_{n,i}(\mathbf{p}_{n,i}^*(p_{n,i}))}{\partial p_{n,i}^{(2)}}\}$.

Finally, by substituting the expressions of the first-order partial derivatives into the the three cases, we can almost get the necessary conditions in (6.8) by some simple algebra manipulations. Slight modifications are needed noticing that Case (i) can possibly occur with $\mathbf{p}_{n,i}^*(p_{n,i}) = (p_{n,i}, 0)^T$ and $\mathbf{p}_{n,i}^*(p_{n,i}) = (0, p_{n,i})^T$ besides $\mathbf{p}_{n,i}^*(p_{n,i}) \succ \mathbf{0}_2$.

E.3 Proof of Theorem 6.4

We first prove $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ is strictly quasiconcave in $\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|}$. The superlevel set of $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ for $\omega \in \mathbb{R}$ is

$$\mathcal{G}_\omega = \left\{\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|} \mid \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n) \geq \omega\right\} = \left\{\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|} \;\middle|\; \omega\left(4p_c^D|\mathcal{H}_n| + 4p_c^S + \rho\sum_{i\in\mathcal{H}_n} p_{n,i}\right) - \sum_{i\in\mathcal{H}_n} C_{n,i}(p_{n,i}) \leq 0\right\}.$$

According to [76], $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ is strictly quasiconcave in $\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|}$ if $\mathcal{G}_\omega$ is strictly convex for any real number ω. Apparently, $\sum_{i\in\mathcal{H}_n} C_{n,i}(p_{n,i})$ is strictly concave in $\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|}$ according to Theorem 6.2 while $\rho\sum_{i\in\mathcal{H}_n} p_{n,i}$ is affine in $\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|}$. Then, $\mathcal{G}_\omega$ is indeed a strictly convex

set for $\omega \in \mathbb{R}$. Therefore, $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ is strictly quasiconcave in $\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|}$. Moreover, the strict quasiconcavity implies for $\mathbf{0}_{|\mathcal{H}_n|} \preceq \mathbf{p}_n \prec \widetilde{\mathbf{p}}_n$ and $0 < \lambda < 1$, we always have

$$\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\lambda \mathbf{p}_n + (1-\lambda)\widetilde{\mathbf{p}}_n) > \min\left\{\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n), \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\mathbf{p}}_n)\right\}. \tag{E.10}$$

Apparently, $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ is continuously differentiable w.r.t $\mathbf{p}_n$. To prove the strict pseudo-concavity of $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$, we need to show that $(\widetilde{\mathbf{p}}_n - \mathbf{p}_n)^T \nabla_{\mathbf{p}_n} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n) < 0$ implies $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\mathbf{p}}_n) < \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ for any $\mathbf{0}_{|\mathcal{H}_n|} \preceq \mathbf{p}_n \prec \widetilde{\mathbf{p}}_n$ [97]. Suppose that there exist $\mathbf{0}_{|\mathcal{H}_n|} \preceq \mathbf{p}_n \prec \widetilde{\mathbf{p}}_n$ such that $(\widetilde{\mathbf{p}}_n - \mathbf{p}_n)^T \nabla_{\mathbf{p}_n} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n) < 0$ and $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\mathbf{p}}_n) \geq \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$. Using Taylor expansion, it is easy to prove there exist $\widetilde{\lambda} \in (0,1)$ such that $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\lambda}\mathbf{p}_n + (1-\widetilde{\lambda})\widetilde{\mathbf{p}}_n) < \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$, based on $(\widetilde{\mathbf{p}}_n - \mathbf{p}_n)^T \nabla_{\mathbf{p}_n} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n) < 0$. Then, $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\lambda}\mathbf{p}_n + (1-\widetilde{\lambda})\widetilde{\mathbf{p}}_n) < \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n) \leq \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\mathbf{p}}_n)$. That is,

$$\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\lambda}\mathbf{p}_n + (1-\widetilde{\lambda})\widetilde{\mathbf{p}}_n) < \min\left\{\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n), \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\mathbf{p}}_n)\right\}, \tag{E.11}$$

which contradicts with (E.10). Thus, $(\widetilde{\mathbf{p}}_n - \mathbf{p}_n)^T \nabla_{\mathbf{p}_n} \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n) < 0$ does imply $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\widetilde{\mathbf{p}}_n) < \widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ for any $\mathbf{0}_{|\mathcal{H}_n|} \preceq \mathbf{p}_n \prec \widetilde{\mathbf{p}}_n$. Therefore, $\widehat{\mathcal{E}}_n^{(\mathcal{H}_n)}(\mathbf{p}_n)$ is strictly pseudoconcave in $\mathbf{p}_n \succeq \mathbf{0}_{|\mathcal{H}_n|}$.

E.4 Proof of Theorem 6.6

The proof is based on the two equivalent forms of problem (6.5) in (6.14) and (6.15). Following the notations in (6.14) and (6.15) but simplifying the superscript of $\mathcal{H}_n^{(k)}$ as k, let $\mathbf{p}_n^{(k)} \triangleq (p_{n,i}^{(k)})_{i \in \mathcal{H}_n^{(k)}}$ and $p_n^{(k)}$ denote the optimal arguments for problems (6.14) and (6.15) when $\mathcal{H}_n = \mathcal{H}_n^{(k)}$, respectively. Then, we have

$$\underbrace{\widehat{\mathcal{E}}_n^{(\mathcal{H}_n^{(k)})}}_{\triangleq \widehat{\mathcal{E}}_n^{(k)}} = \frac{\sum_{i \in \mathcal{H}_n^{(k)}} C_{n,i}(p_{n,i}^{(k)})}{4p_c^D k + 4p_c^S + \rho \sum_{i \in \mathcal{H}_n^{(k)}} p_{n,i}^{(k)}} = \frac{C_n^{(k)}(p_n^{(k)})}{4p_c^D k + 4p_c^S + \rho p_n^{(k)}}. \tag{E.12}$$

For brevity, simplify the notation of $i_n^{(k)}$ as i_k while simplifying $\mathcal{H}_n^{(k)} \bigcup \{i_n^{(k')}\}$ as $k \star i_{k'}$ for $k' \geq k+1$. Because of the definition of $\mathcal{P}_n$, we have $\widehat{\mathcal{E}}_n^{(k+1)} = \max_{k' \geq k+1} \widehat{\mathcal{E}}_n^{(k \star i_{k'})}$. Hence, altogether with the given condition, we have $\widehat{\mathcal{E}}_n^{(k)} > \widehat{\mathcal{E}}_n^{(k \star i_{k'})}$ for $k' \geq k+1$. That is, we have (E.13) for $k' \geq k+1$.

$$\widehat{\mathcal{E}}_n^{(k)} = \frac{C_n^{(k)}(p_n^{(k)})}{4p_c^D k + 4p_c^S + \rho p_n^{(k)}} > \frac{C_n^{(k \star i_{k'})}(p_n^{(k \star i_{k'})})}{4p_c^D (k+1) + 4p_c^S + \rho p_n^{(k \star i_{k'})}} = \frac{C_n^{(k)}(p_n^{(k)}) + [C_n^{(k \star i_{k'})}(p_n^{(k \star i_{k'})}) - C_n^{(k)}(p_n^{(k)})]}{4p_c^D k + 4p_c^S + \rho p_n^{(k)} + [4p_c^D + \rho(p_n^{(k \star i_{k'})} - p_n^{(k)})]}. \tag{E.13}$$

Moreover, from (E.13), we can further infer

$$0 \leq \max_{p_n^{(k)} \leq p_n \leq p_n^{\max}} \frac{C_n^{(k\star i_{k'})}(p_n) - C_n^{(k)}(p_n^{(k)})}{4p_c^D + \rho(p_n - p_n^{(k)})} < \widehat{\mathcal{E}}_n^{(k)}, \tag{E.14}$$

where the first inequality follows from the simple fact that $C_n^{(k\star i_{k'})}(p_n) - C_n^{(k)}(p_n^{(k)}) \geq C_n^{(k\star i_{k'})}(p_n^{(k)}) - C_n^{(k)}(p_n^{(k)}) \geq 0$ over $[p_n^{(k)}, p_n^{\max}]$. The strict inequality in (E.14) has to hold. Otherwise, Eq. (E.13) would contradict with the fact that (i) $p_n^{(k\star i_{k'})}$ is EE-optimal for $\mathcal{H}_n = \mathcal{H}_n^{(k)} \bigcup \{i_n^{(k')}\}$; and (ii) $\frac{a}{b} \leq \frac{a+c}{b+d}$ iff $\frac{a}{b} \leq \frac{c}{d}$, where $a > 0, b > 0, c > 0, d > 0$. That is, there would exist a certain $p_n \in [p_n^{(k)}, p_n^{\max}]$, which was better than $p_n^{(k\star i_{k'})}$ and made $\widehat{\mathcal{E}}_n^{(k)} \leq \widehat{\mathcal{E}}_n^{(k\star i_{k'})}$.

On the other hand, for any $k' \geq k+2$, $\mathcal{H}_n^{(k')} = \mathcal{H}_n^{(k)} \bigcup \{i_{k+1}\} \bigcup \cdots \bigcup \{i_{k'}\}$. Hence, $\widehat{\mathcal{E}}_n^{(k')}$ can be expressed as in (E.15).

$$\widehat{\mathcal{E}}_n^{(k')} = \frac{C_n^{(k')}(p_n^{(k')})}{4p_c^D k' + 4p_c^S + \rho p_n^{(k')}} = \frac{C_n^{(k)}(\sum_{m=1}^{k} p_{n,i_m}^{(k')}) + \sum_{m=k+1}^{k'} [C_{n,i_m}(p_{n,i_m}^{(k')}) + C_n^{(k)}(p_n^{(k)}) - C_n^{(k)}(p_n^{(k)})]}{4p_c^D k + 4p_c^S + \rho \sum_{m=1}^{k} p_{n,i_m}^{(k')} + \sum_{m=k+1}^{k'} [4p_c^D + \rho(p_{n,i_m}^{(k')} + p_n^{(k)} - p_n^{(k)})]}. \tag{E.15}$$

Based on (E.15), we will prove $\widehat{\mathcal{E}}_n^{(k')} < \widehat{\mathcal{E}}_n^{(k)}$ for $k \geq k+2$ in three steps.

Firstly, for any $m \in [k+1, k']$, we have

$$0 \leq \frac{\left(C_{n,i_m}(p_{n,i_m}^{(k')}) + C_n^{(k)}(p_n^{(k)})\right) - C_n^{(k)}(p_n^{(k)})}{4p_c^D + \rho\left(\left(p_{n,i_m}^{(k')} + p_n^{(k)}\right) - p_n^{(k)}\right)} \leq \max_{p_n^{(k)} \leq p_n \leq p_n^{\max}} \frac{C_n^{(k\star i_m)}(p_n) - C_n^{(k)}(p_n^{(k)})}{4p_c^D + \rho(p_n - p_n^{(k)})} < \widehat{\mathcal{E}}_n^{(k)}, \tag{E.16}$$

where the strict inequality is due to (E.14). Therefore, we further have

$$0 \leq \frac{\sum_{m=k+1}^{k'} [C_{n,i_m}(p_{n,i_m}^{(k')}) + C_n^{(k)}(p_n^{(k)}) - C_n^{(k)}(p_n^{(k)})]}{\sum_{m=k+1}^{k'} [4p_c^D + \rho(p_{n,i_m}^{(k')} + p_n^{(k)} - p_n^{(k)})]} < \widehat{\mathcal{E}}_n^{(k)}, \tag{E.17}$$

where the strict inequality is based on (E.16) and the simple fact that $\frac{a}{b} > \frac{\sum_m a_m}{\sum_m b_m}$ if $a_m \geq 0, b_m > 0$ and $\frac{a}{b} > \frac{a_m}{b_m}$.

Secondly, we apparently have

$$0 \leq \frac{C_n^{(k)}(\sum_{m=1}^{k} p_{n,i_m}^{(k')})}{4p_c^D k + 4p_c^S + \rho \sum_{m=1}^{k} p_{n,i_m}^{(k')}} \leq \max_{0 \leq p_n \leq p_n^{\max}} \frac{C_n^{(k)}(p_n)}{4p_c^D k + 4p_c^S + \rho p_n} = \widehat{\mathcal{E}}_n^{(k)}. \tag{E.18}$$

Finally, from (E.15), (E.17), and (E.18), we have $\widehat{\mathcal{E}}_n^{(k')} < \widehat{\mathcal{E}}_n^{(k)}$ for $k \geq k+2$.

REFERENCES

[1] "Rising to meet the 1000x mobile data challenge," QUALCOMM, Tech. Rep., Sept. 2013. [Online]. Available: http://www.qualcomm.com/media/documents/rising-meet-1000x-mobile-data-challenge

[2] T. Edler and S. Lundberg, "Energy efficiency enhancements in radio access networks," in *Erricsson Review*, 2004.

[3] G. Miao, N. Himayat, G. Y. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: a survey," *Wiley J. Wireless Commun. Mobile Comput.*, vol. 9, no. 4, pp. 529–542, Apr. 2009.

[4] R. Kumar and L. Mieritz, "Conceptualizing 'green' IT and data center power and cooling issues," Gartner, Research Paper G00150322, Sept. 2007.

[5] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental tradeoffs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, June 2011.

[6] Y. Li and G. Stüber, *Orthogonal Frequency Division Multiplexing for Wireless Communications*. Springer-Verlag, 2006.

[7] G. Stüber, J. Barry, S. McLaughlin, Y. Li, M. A. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.

[8] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "Ofdm and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[9] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS long term evolution: From theory to practice*. New York: John Wiley & Sons, 2009.

[10] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: Understanding broadband wireless networking*. Pearson Education, 2007.

[11] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[12] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[13] J. Ma, G. Y. Li, and B. H. Juang, "Signal processing in cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.

[14] L. Lu, X. Zhou, U. Onunkwo, and G. Y. Li, "Ten years of research in spectrum sensing and sharing in cognitive radio," *EURASIP J. Wireless Commun. and Netw. 2012*, 2012.

[15] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[16] M. Chen and A. Yener, "Power allocation for F/TDMA multiuser two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 546–551, 2010.

[17] M. Zhou, Q. Cui, R. Jantti, and X. Tao, "Energy-efficient relay selection and power allocation for two-way relay channel with analog network coding," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 816–819, June 2012.

[18] C. Sun and C. Yang, "Energy-efficient hybrid one- and two-way relay transmission," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 3737–3751, 2013.

[19] X. Wang, G. Giannakis, and A. Marques, "A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks," *IEEE Proc.*, vol. 95, no. 12, pp. 2410–2431, Dec. 2007.

[20] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[21] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Mimimum energy to send k bits with and without feedback," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'10)*, Austin, USA, Jun. 2010, pp. 221–225.

[22] M. C. Gursoy, "On the capacity and energy efficiency of training-based transmissions over fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4543–4567, Oct. 2009.

[23] V. Chandar, A. Tchamkerten, and D. Tse, "Asynchronous capacity per unit cost," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'10)*, Austin, USA, Jun. 2010, pp. 280–284.

[24] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sept. 2005.

[25] ——, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.

[26] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 545–554, 2010.

[27] G. Miao, N. Himayat, G. Y. Li, and D. Bormann, "Energy efficient design in wireless OFDMA," in *Proc. IEEE Int. Commun. Conf. (ICC'08)*, Beijing, China, May 2008.

[28] C. Isheden and G.P.Fettweis, "Energy-efficient multi-carrier link adaptation with sum rate-dependent circuit power," in *Proc. IEEE Global Telecommun. Conf. (Globecom'10)*, Miami, FL, US, Dec. 2010.

[29] C. Bae and W. E. Stark, "End-to-end energy-bandwidth tradeoff in multihop wireless networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4051–4066, Sept. 2009.

[30] ——, "On the energy-bandwidth tradeoff for AWGN relay channels," in *Proc. IEEE Military Commun. Conf. (MILCOM'09)*, Boston, MA, US, Oct. 2009.

[31] C. He, B. Sheng, P. Zhu, and X. You, "Energy efficiency and spectral efficiency tradeoff in downlink distributed antenna systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 3, pp. 153–156, June 2012.

[32] C. He, B. Sheng, P. Zhu, X. You, and G. Y. Li, "Energy- and spectral-efficiency tradeoff for distributed antenna systems with proportional fairness," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 894–902, May 2013.

[33] F. Héiot, M. A. Imran, and R. Tafazolli, "On the energy efficiency-spectral efficiency trade-off over the MIMO rayleigh fading channel," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1345–1356, May 2012.

[34] L. Deng, Y. Rui, P. Cheng, J. Zhang, Q. Zhang, and M. Li, "A unified energy efficiency and spectral efficiency tradeoff metric in wireless networks," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 55–58, Jan. 2013.

[35] J. Joung, C. K. Ho, and S. Sun, "Spectral efficiency and energy efficiency of OFDM systems: Impact of power amplifiers and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 208–220, Feb. 2014.

[36] M. Bohge, J. Gross, M. Meyer, and A. Wolisz, "Dynamic resource allocation in OFDM systems: an overview of cross-layer optimization principles and techniques," *IEEE Network Mag.*, vol. 21, no. 1, pp. 53–59, Feb. 2007.

[37] R. S. Prabhu and B. Daneshrad, "An energy-efficient water-filling algorithm for OFDM systems," in *Proc. IEEE Int. Conf. Commun. (ICC'10)*, Cape Town, South Africa, May 2010.

[38] G. Miao, N. Himayat, G. Y. Li, and S. Talwar, "Low-complexity energy-efficient OFDMA," in *Proc. IEEE Int. Conf. Commun. (ICC'09)*, Dresden, Germany, Jun. 2009.

[39] L. Dai, J. Wang, Z. Wang, P. Tsiaflakis, and M. Moonen, "Spectrum- and energy-efficient ofdm based on simultaneous multi-channel reconstruction," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6047–6059, Dec. 2013.

[40] C. He, G. Li, F.-C. Zheng, and X. You, "Energy-efficient resource allocation in OFDM systems with distributed antennas," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1223–1231, Mar. 2014.

[41] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Commun.*, vol. 11, no. 10, pp. 3618–3631, Oct. 2012.

[42] ——, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Commun.*, vol. 11, no. 9, pp. 3292–3304, Sept. 2012.

[43] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 630–643, July 2003.

[44] J. Tang and X. Zhang, "Quality-of-service driven power and rate adapation for multi-channel communications over wireless links," *IEEE Trans. Commun.*, vol. 6, no. 12, pp. 4349–4360, Dec. 2007.

[45] D. Qiao, M. C. Gursoy, and S. Veipasalar, "Energy efficiency in the low-SNR regime under queueing constraints and channel uncertainty," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 2006–2017, July 2011.

[46] H. Zhang, Y. Ma, D. Yuan, and H.-H. Chen, "Quality-of-service driven power and sub-carrier allocation policy for vehicular communication networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 197–206, Jan. 2011.

[47] M. C. Gursory, D. Qiao, and S. Velipasalar, "Analysis of energy efficiency in fading channels under QoS constraints," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4252–4263, August 2009.

[48] E. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. Siris, "Spectrum assignment in cognitive radio: A comprehensive survey," *IEEE Surveys & Tutorials*, Accepted for publication.

[49] Y. Pei, Y.-C. Liang, K. Teh, and K. H. Li, "Energy-efficient design of sequential channel sensing in cognitive radio networks: Optimal sensing strategy, power allocation, and sensing order," *IEEE J. Sel. Area Commun.*, vol. 29, no. 8, pp. 1648–1659, Sept. 2011.

[50] J. Mao, G. Xie, J. Gao, and Y. Liu, "Energy efficiency optimization for OFDM-based cognitive radio systems: A water-filling factor aided search method," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2366–2375, May 2013.

[51] S. Wang, M. Ge, and W. Zhao, "Energy-efficient resource allocation for OFDM-based cognitive radio networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3181–3191, Aug. 2013.

[52] S. Bayhan and F. Alagoz, "Scheduling in centralized cognitive radio networks for energy efficiency," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 582–595, Feb. 2013.

[53] R. Z. Ho, Chin Keong and Y.-C. Liang, "Two-way relaying over OFDM: Optimized tone permutation and power allocation," in *IEEE Proc. Int. Conf. Commun. (ICC'08)*, Beijing, China, May 2008.

[54] Y.-U. Jang, E.-R. Jeong, and Y. H. Lee, "A two-step approach to power allocation for OFDM signals over two-way amplify-and-forward relay," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2426–2430, 2010.

[55] K. Jitvanichphaibool, R. Zhang, and Y.-C. Liang, "Optimal resource allocation for two-way relay-assisted OFDMA," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3311–3321, 2009.

[56] H. Zhang, Y. Liu, and M. Tao, "Resource allocation with subcarrier pairing in OFDMA two-way relay networks," *IEEE Wireless Commun. Lett.*, vol. 1, no. 2, pp. 61–64, Apr. 2012.

[57] C. Sun, Y. Cen, and C. Yang, "Energy efficient OFDM relay systems," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1797–1809, 2013.

[58] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[59] Z. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 56, no. 10, pp. 57–73, Feb. 2008.

[60] P. Tsiaflakis, I. Necoara, J. A. K. Suykens, and M. Moonen, "Improved dual decompostion based optimization for DSL dynamic spectrum management," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2230–2245, Apr. 2010.

[61] O. Arnold, F. Richter, G. P. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Proc. of 19th Future Network & MobileSummit 2010*, Florence, Italy, June 2010.

[62] X. Wang, G. B. Giannakis, and A. G. Marques, "A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks," *Proc. IEEE*, vol. 95, no. 12, pp. 2410–2431, Dec. 2007.

[63] R. K. Sundaram, *A First Course in Optimization*. Cambridge University Press, 1996.

[64] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'06)*, Seattle, USA, July 2006.

[65] X. Zhou, G. Y. Li, D. Li, D. Wang, and A. C. K. Soong, "Probabilistic resource allocation for opportunistic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2870–2879, Sept. 2010.

[66] M. K. Awad, V. Mahinthan, M. Mehrjoo, X. Shen, and J. W. Mark, "A dual-decompostion-based resource allocation for OFDMA networks with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2394–2403, June 2010.

[67] N. Mokari, M. R. Javan, and K. Navaie, "Cross-layer resource allocation in OFDMA systems for heterogeneous traffic with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 1011–1017, Feb. 2010.

[68] I. C. Wong and B. L. Evans, "Optimal resource allocation in the OFDMA downlink with imperfect channel knowledge," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 232–241, Jan. 2009.

[69] C. Y. Wong, R. S. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[70] W. Yu and J. M. Coiffi, "FDMA capacity of Gaussian multiple-access channels with ISI," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 102–111, Jan. 2002.

[71] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subcarrier allocation," in *Proc. IEEE Vehicular Technology Conf. (VTC'00)*, Tokyo, Japan, May 2000, pp. 1085–1089.

[72] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.

[73] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1150–1158, Nov. 2003.

[74] Y. Lin, T. Chiu, and Y. T. Su, "Optimal and near-optimal resource allocation algorithms for OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4066–4077, Aug. 2009.

[75] *Technical specification group GSM/EDGE radio access network radio transmission and reception (release 1999)*, 3GPP TS 05.05 V8.15.0, Aug. 2001.

[76] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[77] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'06)*, Seattle, USA, July 2006.

[78] Z. Mao and X. Wang, "Efficient optimal and suboptimal radio resource allocation in OFDMA system," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 440–445, Feb. 2008.

[79] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 3874–3886, Nov. 2011.

[80] J. Tang and X. Zhang, "Cross-layer-model based adaptive resource allocation for statistical QoS guarantees in mobile wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2318–2328, June 2008.

[81] L. Liu and J.-F. Chamberland, "On the effective capacities of multipleantenna Gaussian channels," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT'08)*, Toronto, Canada, July 2008.

[82] D. S. W. Hui, V. K. N. Lau, and H. L. Wong, "Cross-layer desing for OFDMA wireless systems with heterogeneous delay," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2872–2880, Aug. 2007.

[83] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Commun. Mag.*, vol. 18, no. 6, 2011.

[84] J. Tang and X. Zhang, "Cross-layer modeling for quality of service guarantees over wireless links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4504–4512, Dec. 2007.

[85] C. Xiong, G. Y. Li, S.-Q. Zhang, Y. Chen, and S.-G. Xu, "Energy-efficient resource allocation in OFDMA networks," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3767–3778, Dec 2012.

[86] H. P. Benson, "Using concave envelopes to globally solve the nonlinear sum of ratios problem," *Journal of Global Optimization 22*, pp. 343–364, 2002.

[87] ——, "Global optimization algorithm for the nonlinear sum of ratios problem," *Journal of Optimization Theory and Applications*, vol. 112, no. 1, pp. 1–229, Jan. 2002.

[88] L. Gao and J. Shi, "A numerical sutdy on B&B algorithms for solving sum-of-ratios problem," *T. H. Kim and Adeli (Eds.): AST/UCMA/ISA/ACN 2010, LNCS 6059*, pp. 356–362, 2010.

[89] R. Horst and H. Tuy, *Global Optimization: Deterministic approach*. Berlin: Springer Verlag, 1993.

[90] Z. Xu, C. Yang, G. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient configuration of spatial and frequency resources in MIMO-OFDMA systems," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 564–575, 2013.

[91] G. Miao, N. Himayat, G. Li, and S. Talwar, "Low-complexity energy-efficient scheduling for uplink OFDMA," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 112–120, 2012.

[92] C. He, B. Sheng, P. Zhu, X. You, and G. Li, "Energy- and spectral-efficiency tradeoff for distributed antenna systems with proportional fairness," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 894–902, 2013.

[93] M. Pischella and D. Le Ruyet, "Optimal power allocation for the two-way relay channel with data rate fairness," *IEEE Commun. Lett.*, vol. 15, no. 9, pp. 959–961, Sept. 2011.

[94] S. Schaible and J. Shi, "Fractional programming: The sum-of-ratios case," *Optimization Methods and Software*, vol. 18, no. 2, pp. 219–229, 2003.

[95] T. J. Ypma, "Historical development of the Newton-Raphson method," *SIAM review*, vol. 37, no. 4, pp. 531–551, Dec. 1995.

[96] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*, 2nd ed. New York: Wiley-interscience, 2001.

[97] O. L. MANGASARIAN, "Pseudo-convex functions," *SIAM J. Control Optim.*, vol. Ser. A, no. 3, pp. 281–290, 1965.